```
                                    GenCore version 5.1.6
                                    Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:               October 26, 2005, 13:04:49 ; Search time 165 Seconds
                                                   (without alignments)
                                    583.657 Million cell updates/sec

Title:               US-09-847-513A-7
Perfect score:       1301
Sequence:            1 MKLLILGSVIALPTPAAGG.........NKILFGLIIWNVAVKESSNA 249

Scoring table:       BLOSUM62
                     Gapop 10.0 , Gapext 0.5
Searched:            2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:          2105600

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 78%
                 Listing first 45 summaries

Database :           A_Geneseq_16Dec04:*
                     1: geneseqp1980s:*
                     2: geneseqp1990s:*
                     3: geneseqp2000s:*
                     4: geneseqp2001s:*
                     5: geneseqp2002s:*
                     6: geneseqp2003as:*
                     7: geneseqp2003bs:*
                     8: geneseqp2004s:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.
```

SUMMARIES

```
                   %
Result          Query
No.    Score    Match Length DB  ID            Description
------ -------- ----- ------ --- ----------    -----------
  1    1014.5   78.0   251   5   AAU81215      Proteorho
  2    1014.5   78.0   251   5   AAU81215      Proteorho
  3    1014.5   78.0   251   5   AAU81193      Proteorho
  4    1014.5   78.0   252   8   AAU81216      Proteorho
  5    1014.5   78.0   252   8   ADQ82560      Wild type
  6    1014.5   78.0   252   8   ADQ82562      Wild type
  7    1013.5   77.9   249   8   ADQ82516      Wild type
  8    1012.5   77.8   252   8   ADQ82634      Wild type
  9    1012.5   77.8   251   8   AAU81208      Proteorho
 10    1012.5   77.8   252   8   ADQ82546      Wild type
 11    1011.5   77.7   251   8   AAU81214      Proteorho
 12    1011.5   77.7   252   8   ADQ82558      Wild type
 13    1010.5   77.7   251   5   AAU81209      Proteorho
 14    1009.5   77.6   252   8   ADQ82548      Wild type
 15    1009.5   77.6   251   5   AAU81211      Proteorho
 16    1008.5   77.5   252   8   ADQ82552      Wild type
 17    1008.5   77.5   251   5   AAU81213      Proteorho
 18    1003.5   77.1   252   8   ADQ82556      Wild type
 19    1003.5   77.1   251   5   AAU81192      Proteorho
 20    999.5    76.8   252   8   ADQ82506      Wild type
 21    999.5    76.8   252   8   ADQ82688      Mutant ma
 22    997.5    76.7   252   8   ADQ82686      Mutant ma
 23    996.5    76.6   252   8   ADQ82684      Mutant ma
 24    995.5    76.5   252   8   ADQ82680      Mutant ma
 25    995.5    76.5   252   8   ADQ82678      Mutant ma
 26    994.5    76.4   252   8   ADQ82674      Mutant ma
 27    993.5    76.4   252   8   ADQ82682      Mutant ma
 28    794      61.0   254   8   ADQ82662      Wild type
 29    794      61.0   254   8   ADQ82660      Wild type
 30    776      59.6   254   8   ADQ82658      Wild type
 31    764      58.7   254   8   ADQ82656      Wild type
 32    737      56.6   230   8   ADQ82666      Wild type
 33    737      56.6   259   7   ADF83525      H. salina
 34    177      13.6   259   7   ADF83525      H. salina
 35    170      13.1   262   2   AAR60691      Fragment
 36    169.5    13.1   248   2   AAR26419      Bacterior
 37    169.5    13.0   248   2   AAR60692      Mature H.
 38    169.5    13.0   248   2   AAW53145      Halobacte
 39    168.5    13.0   248   3   AAO17873      Proton tr
 40    165      12.7   214   5   AAB11008      Halobacte
 41    165      12.7   222   5   AAE17868      Bacterior
 42    135.5    10.4   250   2   AAW42105      Amino aci
 43    126      9.7    370   4   AAU43530      Propionib
 44    126      9.7    370   6   ABM40049      Propionib
 45    117.5    9.0    233   5   AAE17866      Halorhodo
```

ALIGNMENTS

```
RESULT 1
AAU81215
ID   AAU81215 standard; protein; 251 AA.
XX
AC   AAU81215;
XX
DT   07-AUG-2003  (revised)
DT   09-APR-2002  (first entry)
XX
DE   Proteorhodopsin from clone PALE6.
XX
KW   Proteorhodopsin; light-driven energy generator; targeted drug delivery;
KW   biocatalytic reactor; fuel cell; nano-machine; molecular switching;
KW   data storage; membrane potential; halophilic arachaea.
XX
OS   Eubacteria.
XX
PN   WO200183701-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014994.
XX
PR   03-MAY-2000; 2000US-0201602P.
XX
PA   (MONT-) MONTEREY BAY AQUARIUM RES INST.
XX
PI   Delong EF,  Beja O;
XX
PR   WPI; 2002-114151/15.
DR   N-PSDB; ABK24154.
XX
PT   New proteorhodopsin gene for use in a light-driven energy generator,
PT   comprises an isolated DNA sequence encoding a proteorhodopsin protein.
XX
PS   Claim 35; Fig 35; 460pp; English.
XX
CC   The invention describes a proteorhodopsin gene, comprising an isolated
CC   DNA sequence for encoding a proteorhodopsin protein. The gene is useful
CC   in light-driven energy generators. A proteorhodopsin system is useful in
CC   many industrial and technological applications, for use in targeted drug
CC   delivery, has primary and secondary energy generators for biocatalytic
CC   reactors, fuel cells and nano-machines, as well as uses in molecular
CC   switching or data storage devices. Proteorhodopsin is useful for a
CC   process to enhance the potential of recombinant protein
CC   production or convert the light-induced membrane potential into cellular
CC   signals. The system is not restricted to operate in halophilic arachaea
CC   and could therefore be functionally expressed in Escherichia coli and
```

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: October 26, 2005, 13:04:49 ; Search time 165 Seconds
                                     (without alignments)
                                     630.061 Million cell updates/sec

Title:         US-09-847-513A-7
Perfect Score: 1301
Sequence:      1 MKLILILGSVIALPTFAAGG..........NKILFGLIIWNVAVKESSNA 249

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5
Searched:      1862994 seqs, 41751061.9 residues

Total number of hits satisfying chosen parameters:    1862902
Listing first 45 summaries

Database :     Published Applications AA.*
1:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
4:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
5:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6:  /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
7:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
9:  /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 78%

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result                  %
No.    Score  Query Match Length  DB  DB ID                    Description
-----  -----  -----------  ------  --  ----------------------  -----------
1      1014.5  10          251     US-09-847-513A-19           Sequence 19, Appl
2      1014.5  10          251     US-09-847-513A-63           Sequence 63, Appl
3      1014.5  10          251     US-09-847-513A-65           Sequence 65, Appl
4      78.0    17          252     US-10-724-264A-11           Sequence 11, Appl
5      78.0    17          252     US-10-724-264A-55           Sequence 55, Appl
6      1014.5  17          252     US-10-724-264A-57           Sequence 57, Appl
7      77.9    17          249     US-10-724-264A-129          Sequence 129, App
8      1012.5  10          17      US-09-847-513A-49           Sequence 49, Appl
9      1012.5  17          252     US-10-724-264A-41           Sequence 41, Appl
10     1011.5  10          251     US-09-847-513A-61           Sequence 61, Appl
11     1011.5  17          252     US-10-724-264A-53           Sequence 53, Appl
12     1010.5  10          251     US-09-847-513A-51           Sequence 51, Appl
13     1010.5  17          251     US-10-724-264A-43           Sequence 43, Appl
14     1009.5  10          251     US-09-847-513A-55           Sequence 55, Appl
15     1009.5  17          252     US-10-724-264A-47           Sequence 47, Appl
16     1008.5  10          251     US-09-847-513A-59           Sequence 59, Appl
17     1008.5  17          251     US-10-724-264A-51           Sequence 51, Appl
18     1003.5  10          252     US-09-847-513A-17           Sequence 17, Appl
19     1003.5  17          251     US-10-724-264A-1            Sequence 1, Appli
20     999.5   17          252     US-10-724-264A-183          Sequence 183, App
21     998.5   17          252     US-10-724-264A-181          Sequence 181, App
22     997.5   17          252     US-10-724-264A-179          Sequence 179, App
23     995.5   17          252     US-10-724-264A-173          Sequence 173, App
24     995.5   17          252     US-10-724-264A-175          Sequence 175, App
25     994.5   17          252     US-10-724-264A-169          Sequence 169, App
26     993.5   17          252     US-10-724-264A-177          Sequence 177, App
27     993.5   17          254     US-10-724-264A-155          Sequence 155, App
28     794     17          254     US-10-724-264A-157          Sequence 157, App
29     794     17          254     US-10-724-264A-153          Sequence 153, App
30     776     17          254     US-10-724-264A-151          Sequence 151, App
31     764     17          230     US-10-724-264A-161          Sequence 161, App
32     737     18          259     US-10-510-628-3             Sequence 3, Appli
33     177     13          262     US-10-688-221-2             Sequence 2, Appli
34     177     13          228     US-10-343-318-1             Sequence 1, Appli
35     169.5   13          248     US-10-864-335B-1            Sequence 1, Appli
36     169.5   13          214     US-10-343-318-1             Sequence 18, Appl
37     165     12          222     US-09-905-176-18            Sequence 18, Appl
38     157.5   12          341     US-09-905-176-15            Sequence 15, Appl
39     149     11          242     US-10-424-599-283641        Sequence 283641, A
40     149     11          304     US-10-424-599-174538        Sequence 174538, A
41     146.5   11          233     US-10-425-115-346442        Sequence 346442,
42     119.5   9           233     US-10-767-701-51333         Sequence 51333, A
43     119.5   9           270     US-10-767-701-51333         Sequence 51333,
44     117.5   9           233     US-09-905-176-16            Sequence 16, Appl
45     110     8           332     US-10-369-493-1522          Sequence 1522, Ap

                              ALIGNMENTS

RESULT 1
US-09-847-513A-19
Sequence 19, Application US/09847513A
Sequence 19, Application US20030104375A1
Publication No. US20030104375A1
GENERAL INFORMATION:
  APPLICANT: MBARI
  APPLICANT: DeLong, Edward
  APPLICANT: Beja, Oded
  TITLE OF INVENTION: Light-driven energy generation using proteorhodopsin
  FILE REFERENCE: MBA-101
  CURRENT APPLICATION NUMBER: US/09/847,513A
  CURRENT FILING DATE: 2001-05-01
  PRIOR APPLICATION NUMBER: 60/201,602
  PRIOR FILING DATE: 2000-05-03
  NUMBER OF SEQ ID NOS: 65
  SOFTWARE: Patentin version 3.0
  SEQ ID NO 19
  LENGTH: 251
  TYPE: PRT
  ORGANISM: Naturally occurring gamma proteobacterium
US-09-847-513A-19

Query Match 78.0%;  Score 1014.5;  DB 10;  Length 251;
Best Local Similarity 79.5%;  Pred. No. 1.7e-92;
Matches 198;  Conservative 16;  Mismatches 34;  Indels 1;  Gaps 1;

Qy   2 KLLILGSVIALPTFAAGGGDLDASDYTGVSFWLVTAALLASTVFFFVERDRVSAKKTS 61
         |||||| :|||||:|| ||||| || |||||||||||:|| ::|||||||: ||||||
Db   3 KLLILGSAIALPSFAAAGGDLDISDTVGVSFWLVTAGMLAATVFFFVERDQVSAKWKTS 62

Qy  62 LTVSGLVTGIAFWHYMYMRGVWIETGDSPTVFRYIDWLLTVPLLICEFYLILAAATNVAG 121
         |||||| ||||||||:|||||||:|||:||||||||||||||||  :|:|||||||| 
Db  63 LTVSGLITGIAFWHYLYMRGVWIDTGDTPTVFRYIDWLLTVPLQMVEFYLILAAACTSVAA 122

```
                                    GenCore version 5.1.6

                    Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 26, 2005, 13:38:51 ; Search time 76 Seconds
                                                (without alignments)
                                                1267.149 Million cell updates/sec

Title:          US-09-847-513A-7
Perfect score:  1301
Sequence:       1 MKLLLIGSVIALPTFAAGG.........NKILFGLIIWNVAVKESSNA 249

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:      33

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 78%
                 Maximum Match 97%
                 Listing first 45 summaries

Database :       A_Geneseq_16Dec04:*
                 1:  geneseqp1980s:*
                 2:  geneseqp1990s:*
                 3:  geneseqp2000s:*
                 4:  geneseqp2001s:*
                 5:  geneseqp2002s:*
                 6:  geneseqp2003as:*
                 7:  geneseqp2003bs:*
                 8:  geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.


                                    SUMMARIES

Result                   Query
  No.    Score  Match  Length  DB    ID         Description
------   -----  -----  ------  --    ------     -----------
  1      1260   96.8    250    5     AAU81202   Aau81202 Proteorho
  2      1260   96.8    250    5     AAU81199   Aau81199 Proteorho
  3      1260   96.8    251    8     ADQ82528   Adq82528 Wild type
  4      1259   96.8    250    8     ADQ82624   Adq82624 Wild type
  5      1258   96.7    250    5     AAU81196   Aau81196 Proteorho
  6      1258   96.7    251    8     ADQ82522   Adq82522 Wild type
  7      1258   96.7    251    8     ADQ82523   Adq82523 Wild type
  8      1257   96.6    250    8     ADQ82620   Adq82620 Wild type
  9      1257   96.6    250    8     ADQ82622   Adq82622 Wild type
  10     1256   96.5    250    8     AAU81197   Aau81197 Proteorho
  11     1254   96.4    250    5     ADQ82534   Adq82534 Wild type
  12     1250   96.1    249    8     ADQ82628   Adq82628 Wild type
  13     1224   94.1    250    8     ADQ82632   Adq82632 Wild type
  14     1098.5 84.4    257    8     ADQ82630   Adq82630 Wild type
  15     1084   83.3    258    8     ADQ82648   Adq82648 Wild type
  16     1075.5 82.7    251    8     ADQ82652   Adq82652 Wild type
  17     1075.5 82.7    251    8     ADQ82642   Adq82642 Wild type
  18     1072.5 82.4    251    8     ADQ82650   Adq82650 Wild type
  19     1072.5 82.4    250    8     ADQ82644   Adq82644 Wild type
  20     1071.5 82.4    251    8     ADQ82640   Adq82640 Wild type
  21     1069.5 82.2    251    8     ADQ82654   Adq82654 Wild type
  22     1068.5 82.1    251    8     ADQ82646   Adq82646 Wild type
  23     1066.5 81.9    247    5     AAU81210   Aau81210 Proteorho
  24     1065.5 81.8    251    5     AAU81212   Aau81212 Proteorho
  25     1032.5 79.4    252    8     ADQ82554   Adq82554 Wild type
  26     1032.5 79.4    252    8     ADQ82550   Adq82550 Wild type
  27     1026.5 78.9    251    5     AAU81195   Aau81195 Proteorho
  28     1026.5 78.9    251    8     ADQ82638   Adq82638 Wild type
  29     1026.5 78.9    252    8     ADQ82520   Adq82520 Wild type
  30     1019.5 78.4    252    8     ADQ82664   Adq82664 Wild type
  31     1017.5 78.2    249    8     ADQ82636   Adq82636 Wild type
  32     1016.5 78.1    251    5     AAU81194   Aau81194 Proteorho
  33     1016.5 78.1    252    8     ADQ82518   Adq82518 Wild type
```

```
                                    ALIGNMENTS

RESULT 1
ID      AAU81202 standard; protein; 250 AA.
XX
AC      AAU81202;
XX
DT      07-AUG-2003  (revised)
DT      09-APR-2002  (first entry)
XX
DE      Proteorhodopsin from clone MB40m5.
XX
KW      Proteorhodopsin; light-driven energy generator; targeted drug delivery;
KW      biocatalytic reactor; fuel cell; nano-machine; molecular switching;
KW      data storage; membrane potential; halophilic arachaea.
XX
OS      Eubacteria.
XX
PN      WO200183701-A2.
XX
PD      08-NOV-2001.
XX
PF      02-MAY-2001; 2001WO-US014394.
XX
PR      03-MAY-2000; 2000US-0201602P.
XX
PA      (MONT-) MONTEREY BAY AQUARIUM RES INST.
XX
PI      Delong EF, Beja O;
XX
DR      WPI; 2002-114151/15.
DR      N-PSDB; ABK24141.
XX
PT      New proteorhodopsin gene for use in a light-driven energy generator,
PT      comprises an isolated DNA sequence encoding a proteorhodopsin protein.
XX
PS      Claim 22; Fig 22; 460pp; English.
CC      The invention describes a proteorhodopsin gene, comprising an isolated
CC      DNA sequence for encoding a proteorhodopsin protein. The gene is useful
CC      in light-driven energy generators. A proteorhodopsin system is useful in
CC      many industrial and technological applications, for use in targeted drug
CC      delivery, has primary and secondary energy generators for biocatalytic
CC      reactors, fuel cells and nano-machines, as well as uses in molecular
CC      switching or data storage devices. Proteorhodopsin is useful for a
CC      process to enhance yield or increase the potential of recombinant protein
CC      production or convert the light-induced membrane potential into cellular
CC      signals. The system is not restricted to operate in halophilic arachaea
CC      and could therefore be functionally expressed in Escherichia coli and
CC      other bacteria. It provides for a fast and cheap production method that
CC      allows for mass production of functionally active proteorhodopsin. This
CC      is the amino acid sequence of a proteorhodopsin protein variant,
CC      described in the method of the invention. (Updated on 07-AUG-2003 to
CC      correct OS field.)
XX
SQ      Sequence 250 AA;

Query Match             96.8%;  Score 1260;  DB 5;  Length 250;
Best Local Similarity   97.2%;  Pred. No. 4.2e-133;
Matches  241; Conservative   2; Mismatches   5;  Indels   0;  Gaps   0;
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:            October 26, 2005, 13:38:51 ; Search time 162 Seconds
                                                 (without alignments)
                                                 641.729 Million cell updates/sec

Title:             US-09-847-513A-7
Perfect score:     1301
Sequence:          1 MKLLILIGSVIALPTPAAGG.........NKILFGLIIWNVAVKESSNA 249

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          162994 seqs, 417510619 residues

Total number of hits satisfying chosen parameters:           33

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match .78%
                 Maximum Match 97%
                 Listing first 45 summaries

Database :         Published Applications_AA:*
                1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
               10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
               11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
               12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
               13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
               14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
               15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
               16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
               17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
               18: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
               19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
               20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
               21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
               22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1260 | 96.8 | 250 | 10 | US-09-847-513A-31 | Sequence 31, Appl |
| 2 | 1260 | 96.8 | 250 | 17 | US-09-847-513A-37 | Sequence 37, Appl |
| 3 | 1260 | 96.8 | 251 | 17 | US-10-724-264A-23 | Sequence 23, Appl |
| 4 | 1259 | 96.8 | 250 | 10 | US-10-724-264A-119 | Sequence 119, App |
| 5 | 1258 | 96.7 | 250 | 17 | US-09-847-513A-25 | Sequence 25, Appl |
| 6 | 1258 | 96.7 | 251 | 17 | US-10-724-264A-17 | Sequence 17, Appl |
| 7 | 1258 | 96.7 | 251 | 17 | US-10-724-264A-18 | Sequence 18, Appl |
| 8 | 1257 | 96.7 | 250 | 17 | US-10-724-264A-115 | Sequence 115, App |
| 9 | 1256 | 96.5 | 250 | 17 | US-10-724-264A-117 | Sequence 117, App |
| 10 | 1254 | 96.4 | 250 | 10 | US-09-847-513A-27 | Sequence 27, Appl |
| 11 | 1250 | 96.1 | 249 | 17 | US-10-724-264A-29 | Sequence 29, Appl |

ALIGNMENTS

```
RESULT 1
US-09-847-513A-31
; Sequence 31, Application US/09847513A
; Publication No. US20030104375A1
; GENERAL INFORMATION:
;  APPLICANT: MBARI
;  APPLICANT: DeLong, Edward
;  APPLICANT: Beja, Oded
;  TITLE OF INVENTION: Light-driven energy generation using proteorhodopsin
;  FILE REFERENCE: MBA-101
;  CURRENT APPLICATION NUMBER: US/09/847,513A
;  CURRENT FILING DATE: 2001-05-01
;  PRIOR APPLICATION NUMBER: 60/201,602
;  PRIOR FILING DATE: 2000-05-03
;  NUMBER OF SEQ ID NOS: 65
;  SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 31
;  LENGTH: 250
;  TYPE: PRT
;  ORGANISM: Naturally occurring gamma proteobacterium
US-09-847-513A-31

  Query Match             96.8%;
  Best Local Similarity   97.6%;
  Matches  242;        Conservative    3;

              Score 1260;  DB 10;  Length 250;
              Pred. No. 6.3e-117;
              Mismatches    3;    Indels    0;    Gaps    0;

Qy          2 KLLILIGSVIALPTPAAGGGDLDASDYTGVSFWLVTAALLASTVFFVERDRVSAKWKTS 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 KLLILIGSVIALPTPAAGGGDLDASDYTGVSFWLVTAALLASTVFFVERDRVSAKWKTS 62

Qy         62 LTVSGLVTGIAFWHYMYMRGVWIETGDSPTVPRYIDWLLTVPLLICFYLLLAAATNVAG 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 LTVSGLVTGIAFWHYMYMRGVWIETGDSPTVPRYIDWLLTVPLLICFYLLLAAATNVAG 122

Qy        122 SLFKKLLVGSLVMLVPGYMGEAGIMAAWPAFIIGCLAWVYMIYELWAGEGKSACNTASPA 181
              ||||||||||||||||||||||||:||||||||||||||||||||:|||||||||||:
Db        123 SLFKKLLVGSLVMLVPGYMGERAQIMAAWPAFIIGCLAWVYMIYELYAGEGKSACNTASPS 182

Qy        182 VQSAYNTMMYIIIFGWAIYPVGYFTGYLMGDGGSALNLINLIYNLADFVNKILFGLIIWNV 241
              ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 VQSAYNTMMAIIVFGWAIYPVGYFTGYLMGDGGSALNLINLIYNLADFVNKILLGLIIWNV 242

Qy        242 AVKESSNA 249
              ||||||||
Db        243 AVKESSNA 250
```

| 12 | 1224 | 94.1 | 250 | 17 | US-10-724-264A-123 | Sequence 123, App |
| 13 | 1098.5 | 84.4 | 257 | 17 | US-10-724-264A-127 | Sequence 127, App |
| 14 | 1084 | 83.3 | 258 | 17 | US-10-724-264A-125 | Sequence 125, App |
| 15 | 1075.5 | 82.7 | 251 | 17 | US-10-724-264A-143 | Sequence 143, App |
| 16 | 1075.5 | 82.7 | 251 | 17 | US-10-724-264A-147 | Sequence 147, App |
| 17 | 1072.5 | 82.4 | 250 | 17 | US-10-724-264A-137 | Sequence 137, App |
| 18 | 1071.5 | 82.4 | 250 | 17 | US-10-724-264A-145 | Sequence 145, App |
| 19 | 1069.5 | 82.2 | 251 | 17 | US-10-724-264A-135 | Sequence 135, App |
| 20 | 1068.5 | 82.1 | 251 | 17 | US-10-724-264A-139 | Sequence 139, App |
| 21 | 1066.5 | 82.0 | 251 | 17 | US-10-724-264A-141 | Sequence 141, App |
| 22 | 1065.5 | 81.9 | 247 | 17 | US-09-847-513A-53 | Sequence 53, Appl |
| 23 | 1032.5 | 79.4 | 251 | 10 | US-09-847-513A-57 | Sequence 57, Appl |
| 24 | 1032.5 | 79.4 | 251 | 10 | US-10-724-264A-45 | Sequence 45, Appl |
| 25 | 1032.5 | 79.4 | 252 | 17 | US-10-724-264A-49 | Sequence 49, Appl |
| 26 | 1032.5 | 78.9 | 251 | 17 | US-09-847-513A-23 | Sequence 23, Appl |
| 27 | 1026.5 | 78.9 | 251 | 10 | US-10-724-264A-137 | Sequence 137, App |
| 28 | 1026.5 | 78.9 | 252 | 17 | US-10-724-264A-133 | Sequence 133, App |
| 29 | 1026.5 | 78.9 | 252 | 17 | US-10-724-264A-15 | Sequence 15, Appl |
| 30 | 1019.5 | 78.4 | 249 | 17 | US-10-724-264A-159 | Sequence 159, App |
| 31 | 1017.5 | 78.2 | 251 | 17 | US-10-724-264A-131 | Sequence 131, App |
| 32 | 1016.5 | 78.1 | 251 | 10 | US-09-847-513A-21 | Sequence 21, Appl |
| 33 | 1016.5 | 78.1 | 252 | 17 | US-10-724-264A-13 | Sequence 13, Appl |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:            October 25, 2005, 09:20:28 ; Search time 165 Seconds
                   (without alignments)
                   583.657 Million cell updates/sec

Title:             US-09-847-513A-7
Perfect score:     1301
Sequence:          1 MKLLLIGSVIALPTFAAGG...........NKILPGLIIWNVAVKESSNA 249

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:      2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing:   Minimum Match 0%
                   Maximum Match 100%
                   Listing first 45 summaries

Database :         A_Geneseq_16Dec04:*
                   1: genesecp1908:*
                   2: genesecp1990s:*
                   3: geneseqp2000s:*
                   4: geneseqp2001s:*
                   5: geneseqp2002s:*
                   6: geneseqp2003as:*
                   7: geneseqp2003bs:*
                   8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1301 | 100.0 | 249 | 5 | AAU81187 | Aau81187 Proteorho |
| 2 | 1301 | 100.0 | 249 | 8 | ADQ82508 | Adq82508 Wild type |
| 3 | 1296 | 99.6 | 250 | 5 | AAU81186 | Aau81186 Proteorho |
| 4 | 1296 | 99.6 | 250 | 8 | ADQ82592 | Adq82592 Wild type |
| 5 | 1296 | 99.6 | 251 | 8 | ADQ82542 | Adq82542 Wild type |
| 6 | 1294 | 99.5 | 249 | 8 | ADQ82670 | Adq82670 Wild type |
| 7 | 1293 | 99.4 | 249 | 8 | ADQ82672 | Adq82672 Mutant ma |
| 8 | 1293 | 99.4 | 250 | 8 | ADQ82606 | Adq82606 Wild type |
| 9 | 1292 | 99.3 | 249 | 8 | ADQ82668 | Adq82668 Mutant ma |
| 10 | 1292 | 99.3 | 250 | 5 | AAU81189 | Aau81189 Proteorho |
| 11 | 1291 | 99.2 | 250 | 5 | AAU81200 | Aau81200 Proteorho |
| 12 | 1290 | 99.2 | 250 | 8 | ADQ82616 | Adq82616 Wild type |
| 13 | 1289 | 99.1 | 250 | 8 | ADQ82574 | Adq82574 Wild type |
| 14 | 1289 | 99.1 | 250 | 8 | ADQ82570 | Adq82570 Wild type |
| 15 | 1289 | 99.1 | 250 | 8 | ADQ82590 | Adq82590 Wild type |
| 16 | 1289 | 99.1 | 250 | 8 | ADQ82588 | Adq82588 Wild type |
| 17 | 1289 | 99.1 | 250 | 8 | ADQ82582 | Adq82582 Wild type |
| 18 | 1289 | 99.1 | 250 | 8 | ADQ82592 | Adq82592 Wild type |
| 19 | 1288 | 99.0 | 250 | 5 | AAU81207 | Aau81207 Proteorho |
| 20 | 1288 | 99.0 | 251 | 8 | ADQ82544 | Adq82544 Wild type |
| 21 | 1287 | 98.9 | 250 | 8 | ADQ82586 | Adq82586 Wild type |
| 22 | 1287 | 98.9 | 250 | 8 | ADQ82578 | Adq82578 Wild type |
| 23 | 1286 | 98.8 | 250 | 8 | ADQ82566 | Adq82566 Wild type |
| 24 | 1286 | 98.8 | 250 | 8 | ADQ82596 | Adq82596 Wild type |
| 25 | 1285 | 98.8 | 250 | 8 | ADQ82564 | Adq82564 Wild type |
| 26 | 1285 | 98.8 | 250 | 8 | ADQ82594 | Adq82594 Wild type |
| 27 | 1284 | 98.7 | 250 | 5 | AAU81201 | Aau81201 Proteorho |
| 28 | 1284 | 98.7 | 251 | 8 | ADQ82532 | Adq82532 Wild type |
| 29 | 1283 | 98.6 | 250 | 5 | AAU81203 | Aau81203 Proteorho |
| 30 | 1283 | 98.6 | 250 | 8 | ADQ82602 | Adq82602 Wild type |
| 31 | 1283 | 98.6 | 250 | 8 | ADQ82576 | Adq82576 Wild type |
| 32 | 1283 | 98.6 | 251 | 8 | ADQ82536 | Adq82536 Wild type |
| 33 | 1282 | 98.5 | 250 | 8 | ADQ82600 | Adq82600 Wild type |
| 34 | 1282 | 98.5 | 250 | 8 | ADQ82568 | Adq82568 Wild type |
| 35 | 1282 | 98.5 | 250 | 8 | ADQ82580 | Adq82580 Wild type |
| 36 | 1281 | 98.5 | 249 | 8 | ADQ82530 | Adq82530 Wild type |
| 37 | 1280 | 98.4 | 250 | 8 | ADQ82584 | Adq82584 Wild type |
| 38 | 1280 | 98.4 | 250 | 8 | ADQ82598 | Adq82598 Wild type |
| 39 | 1279 | 98.3 | 250 | 8 | ADQ82614 | Adq82614 Wild type |
| 40 | 1277 | 98.2 | 250 | 8 | ADQ82618 | Adq82618 Wild type |
| 41 | 1277 | 98.2 | 250 | 8 | ADQ82608 | Adq82608 Wild type |
| 42 | 1275 | 98.0 | 250 | 8 | ADQ82572 | Adq82572 Wild type |
| 43 | 1274 | 97.9 | 250 | 5 | AAU81204 | Aau81204 Proteorho |
| 44 | 1274 | 97.9 | 251 | 8 | ADQ82538 | Adq82538 Wild type |
| 45 | 1273 | 97.8 | 249 | 8 | ADQ82610 | Adq82610 Wild type |

ALIGNMENTS

RESULT 1
AAU81187

ID   AAU81187 standard; protein; 249 AA.
XX
AC   AAU81187;
XX
DT   07-AUG-2003  (revised)
DT   09-APR-2002  (first entry)
XX
DE   Proteorhodopsin from clone EBAC31A8 #2.
XX
KW   Proteorhodopsin; light-driven energy generator; targeted drug delivery;
KW   biocatalytic reactor; fuel cell; nano-machine; molecular switching;
KW   data storage; membrane potential; halophilic archaea.
XX
OS   Eubacteria.
XX
PN   WO200183701-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014394.
XX
PR   03-MAY-2000; 2000US-0201602P.
XX
PA   (MONT-) MONTEREY BAY AQUARIUM RES INST.
XX
PI   Delong EF, Beja O;
XX
PT   New proteorhodopsin gene for use in a light-driven energy generator,
PT   comprises an isolated DNA sequence encoding a proteorhodopsin protein.
XX
PS   Disclosure; Fig 5; 460pp; English.
XX
DR   WPI; 2002-114151/15.
DR   N-PSDB; ABK24126.
XX
CC   The invention describes a proteorhodopsin gene, comprising an isolated
CC   DNA sequence for encoding a proteorhodopsin protein. The gene is useful
CC   in light-driven energy generators. A proteorhodopsin system is useful in
CC   many industrial and technological applications, for use in targeted drug
CC   delivery, has primary and secondary energy generators for biocatalytic
CC   reactors, fuel cells and nano-machines, as well as uses in molecular
CC   switching or data storage devices. Proteorhodopsin is useful for a
CC   process to enhance yield or increase the potential of recombinant protein
CC   production or convert the light-induced membrane potential into cellular
CC   signals. The system is not restricted to operate in halophilic archaea
CC   and could therefore be functionally expressed in Escherichia coli and

```
                        GenCore version 5.1.6

            Copyright (c) 1993 - 2005 Compugen Ltd.


OM protein - protein search, using sw model

Run on:        October 25, 2005, 09:23:31 ; Search time 169 Seconds
                                            (without alignments)
                                            754.484 Million cell updates/sec

Title:     US-09-847-513A-7
Perfect score:  1301
Sequence:  1 MKLLLILGSVIALPTFAGG..........NKILFGLIIWNVAVKESSNA 249

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:    1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Uniprot_03:*
         1:   uniprot_sprot:*
         2:   uniprot_trembl:*

         Pred. No. is the number of results predicted by chance to have a
         score greater than or equal to the score of the result being printed,
         and is derived by analysis of the total score distribution.


                              SUMMARIES

Result                 %        Query
 No.   *Score   Match Length  DB   ID          Description
-----  ------   ------------  --   ----------  ------------
  1    1301     100.0   249   1    PRRG_PRB01  PRRG_PRB01
  2    1296      99.6   250   2    Q63JG7      uncultured
  3    1296      99.6   250   1    Q6J4G7      gamma-prote
  4    1293      99.4   250   2    Q9AFE6      uncultured
  5    1290      99.2   250   2    Q9AFE6      uncultured
  6    1289      99.1   250   2    Q8BD7       uncultured
  7    1289      99.1   250   2    Q84C28      uncultured
  8    1288      99.0   251   2    Q84BE3      uncultured
  9    1287      98.9   251   2    Q9AFE5      uncultured
 10    1287      98.9   250   2    Q84BE4      uncultured
 11    1286      98.8   250   2    Q3UH6       uncultured
 12    1286      98.8   250   2    Q84BE8      uncultured
 13    1285      98.8   250   2    Q84BE9      uncultured
 14    1285      98.8   251   2    Q84BE2      uncultured
 15    1284      98.7   251   2    Q84FF1      uncultured
 16    1283      98.6   250   2    Q84BD9      uncultured
 17    1283      98.6   251   2    Q84BE6      uncultured
 18    1282      98.5   250   2    Q9AFE9      uncultured
 19    1282      98.5   250   2    Q83TG9      uncultured
 20    1281      98.5   249   2    Q84BE0      uncultured
 21    1281      98.5   250   2    Q9AFF2      uncultured
 22    1280      98.4   250   2    Q84BE1      uncultured
 23    1280      98.4   250   2    Q84BE5      uncultured
 24    1279      98.3   250   2    Q84C29      uncultured
 25    1277      98.2   250   2    Q84C27      uncultured
 26    1275      98.0   250   2    Q84BE7      uncultured
 27    1274      97.9   251   2    Q9AFE8      uncultured
 28    1273      97.8   249   2    Q84C31      uncultured
 29    1273      97.8   250   2    Q84C30      uncultured
 30    1272      97.8   252   2    Q9AFF9      uncultured
 31    1270      97.6   251   2    Q84BD8      uncultured
 32    1268      97.5   250   2    Q84C23      uncultured
 33    1267      97.4   249   2    Q84C32      uncultured
 34    1267      97.4   251   2    Q99GB3      uncultured
 35    1262      97.0   252   2    Q9AFG0      uncultured
 36    1260      96.8   251   2    Q9AFF3      uncultured
 37    1259      96.8   250   2    Q84C24      uncultured
 38    1258      96.7   251   2    Q9AFF5      uncultured
 39    1257      96.6   250   2    Q84C26      uncultured
 40    1256      96.5   250   2    Q84C25      uncultured
 41    1254      96.4   251   2    Q9AFF4      uncultured
 42    1252.5    96.3   250   2    Q9AFF0      uncultured
 43    1250      96.1   251   2    Q6J5K4      uncultured
 44    1239      95.2   244   2    Q6J5K5      uncultured
 45    1237      95.1   250   2    Q84C22      uncultured
```

ALIGNMENTS

RESULT 1
PRRG_PRB01
ID    PRRG_PRB01       STANDARD;      PRT;   249 AA.
AC    Q9F7P4;
DT    10-OCT-2003 (Rel. 42, Created)
DT    10-OCT-2003 (Rel. 42, Last sequence update)
DT    29-MAR-2004 (Rel. 43, Last annotation update)
DE    Green-light absorbing proteorhodopsin precursor (GPR).
OS    Gamma-proteobacterium EBAC31A08.
OC    Bacteria; Proteobacteria; Gammaproteobacteria; environmental samples.
OX    NCBI_TaxID=133804;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    MEDLINE=20446260; PubMed=10988064; DOI=10.1126/science.289.5486.1902;
RA    Beja O., Aravind L., Koonin E.V., Suzuki M.T., Hadd A., Nguyen L.P.,
RA    Jovanovich S.B., Gates C.M., Feldman R.A., Spudich J.L., Spudich E.N.,
RA    DeLong E.F.;
RT    "Bacterial rhodopsin: evidence for a new type of phototrophy in the
RT    sea.";
RL    Science 289:1902-1906(2000).
RN    [2]
RP    CHARACTERIZATION OF PROTON RELEASE CYCLE, AND MUTAGENESIS OF CYS-107;
RP    CYS-156 AND CYS-175.
RX    PubMed=11943070; DOI=10.1186/1472-6793-2-5;
RA    Krebs R.A., Alexiev U., Partha R., DeVita A., Braiman M.S.;
RT    "Detection of fast light-activated H+ release and M intermediate
RT    formation from proteorhodopsin.";
RL    BMC Physiol. 2:5-5(2002).
RN    [3]
RP    CHARACTERIZATION OF PHOTOCHEMICAL CYCLE, AND MUTAGENESIS OF ASP-97 AND
RP    GLU-108.
RX    MEDLINE=21965984; PubMed=11969395; DOI=10.1021/bi025563x;
RA    Dioumaev A.K., Brown L.S., Shih J., Spudich E.N., Spudich J.L.,
RA    Lanyi J.K.;
RT    "Proton transfers in the photochemical reaction cycle of
RT    proteorhodopsin.";
RL    Biochemistry 41:5348-5358(2002).
RN    [4]
RP    CHARACTERIZATION OF PHOTOCHEMICAL CYCLE.
RX    MEDLINE=22343728; PubMed=12547799;
RA    Varo G., Brown L.S., Lakatos M., Lanyi J.K.;
RT    "Characterization of the photochemical reaction cycle of
RT    proteorhodopsin.";
RL    Biophys. J. 84:1202-1207(2003).
RN    [5]
RP    COMPARISON WITH PHOTOCHEMICAL CYCLE OF BLUE PROTEORHODOPSIN.
RX    PubMed=12821661; DOI=10.1074/jbc.M305716200;
RA    Wang W.W., Sineshchekov O.A., Spudich E.N., Spudich J.L.;
RT    "Spectroscopic and photochemical characterization of a deep ocean
RT    proteorhodopsin.";
RL    J. Biol. Chem. 278:33985-33991(2003).
RN    [6]
RP    PRELIMINARY CHARACTERIZATION OF PUMP VECTORIALITY.
RX    MEDLINE=22195742; PubMed=12206764; DOI=10.1016/S0022-2836(02)00696-4;

```
                              GenCore version 5.1.6
                         Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October 25, 2005, 09:38:38 ; Search time 39 Seconds
                                             (without alignments)
                                             614.307 Million cell updates/sec

Title:          1301
Perfect score:  1301
Sequence:       1 MKLLLILGSVIALPTFAAGG.........NKILFGLIIWNVAVKESSNA 249

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:      283416

Minimum DB seq length:0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      PIR_79:*
                1: pir1:*
                2: pir2:*
                3: pir3:*
                4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.


                              SUMMARIES

Result       %    Query
No.  Score  Match Length DB ID    Description
---  -----  ----- ------ -- ----- -----------
 1   187.5  14.4  239  2  S55300  sensory rhodopsin
 2   177    13.6  262  1  RAHSB   bacteriorhodopsin
 3   177    13.6  262  2  H84300  bacteriorhodopsin
 4   169.5  13.0  237  1  T44947  bacteriorhodopsin
 5   169.5  13.0  237  2  H84327  bacteriorhodopsin
 6   169.5  13.0  249  4  T44821  sensory rhodopsin
 7   169.5  13.0  259  4  T44820  bacteriorhodopsin
 8   168    12.9  258  2  T50676  bacteriorhodopsin
 9   168    12.9  260  2  A34178  bacteriorhodopsin
10   163.5  12.6  259  2  S14731  bacteriorhodopsin
11   163    12.5  236  2  S55297  bacteriorhodopsin
12   153.5  11.8  250  2  T50677  sensory rhodopsin
13   149.5  11.5  250  2  S51206  cruxrhodopsin-1 -
14   141.5  10.9  209  2  B47686  bacteriorhodopsin-
15   140    10.8  204  2  C47686  bacteriorhodopsin-
16   137.5  10.6  211  2  A47686  bacteriorhodopsin-
17   134    10.3  291  2  A35502  halorhodopsin - Na
18   129.5  10.0  282  2  T43840  halorhodopsin [sim
19   127     9.8  284  2  S29988  halorhodopsin prec
20   118.5   9.1  274  1  A26161  halorhodopsin prec
21   118.5   9.1  274  2  G84178  halorhodopsin prec
22   117.5   9.0  254  2  T43843  sensory rhodopsin
23   116.5   9.0  276  2  T43842  halorhodopsin [imp
24   113.5   8.7  276  2  T48843  halorhodopsin [sim
25   112     8.6  239  2  A56808  halorhodopsin [val
26   109     8.4  332  2  S31848  heat shock protein
27   109     8.4  512  2  G90399  amino acid transpo
28   106     8.1  261  2  AF2201  bacteriorhodopsin
29   105.5   8.1  407  2  AF1716  antiporter protein
30   105     8.1  298  2  A83892  hypothetical prote
31   105     8.1  361  2  C84246  ribose ABC transpo
32   104     8.0  239  1  S09277  sensory rhodopsin
33   104     8.0  239  2  F84318  sensory rhodopsin
34   102.5   7.9  412  2  A83604  probable MFS trans
35   102.5   7.9  426  2  JC5086  polytopic cytoplas
36   102.5   7.9  453  2  H83790  sodium-dependent t
37   102     7.8  292  2  T49829  related to YRO2 pr
38   102     7.8  490  2  S77201  hypothetical prote
39   101.5   7.8  407  2  AC1346  antiporter protein
40   101.5   7.8  525  2  S26022  cytochrome-c oxida
41   101.5   7.8  606  2  S66857  probable membrane
42   101     7.7  297  2  T43838  halorhodopsin [sim
43   100     7.7  820  2  G75251  cytochrome-c oxida
44    99.5   7.6  247  2  S29989  sensory rhodopsin
45    99.5   7.6  315  2  AB3435  protoheme IX farne
```

```
                              ALIGNMENTS

RESULT 1
S55300
sensory rhodopsin II - Natronobacterium pharaonis
N;Alternate names: phoborhodopsin
C;Species: Natronobacterium pharaonis
C;Date: 14-Oct-1995 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
C;Accession: S55300; S55301
R;Seidel, R.; Scharf, B.; Gautel, M.; Kleine, K.; Oesterhelt, D.; Engelhard, M.
Proc. Natl. Acad. Sci. U.S.A. 92, 3036-3040, 1995
A;Title: The primary structure of sensory rhodopsin II: a member of an additional retinal
A;Reference number: S55296; MUID:95224074; PMID:7708770
A;Accession: S55300
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-239 <SEI>
A;Cross-references: UNIPROT:P42196; EMBL:Z35086; NID:g510868; PID:g510870
A;Experimental source: strain SP1
A;Accession: S55301
A;Molecule type: protein
A;Residues: 2-31 <SEW>
A;Experimental source: strain SP1
A;Gene: SRII
A;Superfamily: bacteriorhodopsin
C;Keywords: chromoprotein; photoreceptor; retinal; transmembrane protein
F;8-25/Domain: transmembrane #status predicted <TM1>
F;33-55/Domain: transmembrane #status predicted <TM2>
F;71-90/Domain: transmembrane #status predicted <TM3>
F;99-115/Domain: transmembrane #status predicted <TM4>
F;127-145/Domain: transmembrane #status predicted <TM5>
F;158-181/Domain: transmembrane #status predicted <TM6>
F;193-215/Domain: transmembrane #status predicted <TM7>
F;205/Binding site: retinal (Lys) (covalent) #status predicted

Query Match      14.4%;  Score 187.5;  DB 2;  Length 239;
Best Local Similarity   30.8%;  Pred. No. 5.2e-09;
Matches   69;  Conservative   42;  Mismatches   84;  Indels   29;  Gaps   13;

Qy      33 FWLVTAALLASTV-FFFVERDRVSAKWKTSLTVSGLVTGIAFWHYMYMR-GV-WIETGDS 89
           ||  |  || || |  ||::|    |:   :|  |   |    | | : |   |  |:|
Db       8 FWLGAIGMLVGTLAFAWAGRDAGSGERRYYVTLVG-ISGIAAVAYVVWALGVGVPVAER 66

Qy      90 PTVF--RYIDWLLTVPLLICEFYLILAAATNVAGSLFKLLVGSLVMLVFGYMGEAGIMA 147
           | |   :||||| || | ||:|    |: :||| ||| |  |  :| ||||  | |  |
Db      67 -TVPAPRYIDWLLTTPLLIVFLGLLAGLDSREFGIV----ITLNTVVMLA----GFAGAMV 118

Qy     148 ----AWPAFIIGCLAWVYMIYELWAGECKSACNTASPAVQSAYNTMMYIIIFGWAIYPVG 203
               :|| |  ||| ||:  : :||    |::  |:  | |  |  ::||| || ||| |
Db     119 PGIERYALFGMGAVAFLGLVYYL-VGPMTESASQRSSGIKSLVRLRNLTVILWAIYPF- 176

Qy     204 YFIGYLMGDGGSAL------NLNLIYNLADFVNKILFGLIIWNVA 242
            :| |  |   |      :  || | || |  ||||||||||||||
Db     204 ------------------NKILFGLIIWNVA 242
```

OM protein - protein search, using sw model

```
                GenCore version 5.1.6
                Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:         October 25, 2005, 09:20:28 ; Search time 43 Seconds
                                             (without alignments)
                                             432.270 Million cell updates/sec

Title:          US-09-847-513A-7
Perfect score:  1301
Sequence:       1 MKLILILGSVIALPTFAAGG.........NKILFGLIIWNVAVKESSNA 249

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:    513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   ;
            Issued_Patents_AA:*
            1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
            2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
            3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
            4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
            5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
            6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 177 | 13.6 | 262 | 1 | US-08-313-553-2 | Sequence 2, Appli |
| 2 | 177 | 13.6 | 262 | 3 | US-08-767-993-2 | Sequence 2, Appl |
| 3 | 177 | 13.6 | 559 | 1 | US-08-313-553-15 | Sequence 15, Appl |
| 4 | 177 | 13.6 | 559 | 4 | US-08-767-993-15 | Sequence 15, Appl |
| 5 | 169.5 | 13.0 | 248 | 1 | US-08-313-553-3 | Sequence 3, Appli |
| 6 | 169.5 | 13.0 | 248 | 1 | US-08-313-553-3 | Sequence 3, Appli |
| 7 | 169.5 | 13.0 | 248 | 1 | US-08-767-993-3 | Sequence 3, Appl |
| 8 | 102.5 | 7.9 | 466 | 4 | US-09-252-991A-27479 | Sequence 27479, A |
| 9 | 102 | 7.8 | 405 | 4 | US-09-602-208A-262 | Sequence 262, App |
| 10 | 99 | 7.6 | 415 | 4 | US-09-489-039A-13505 | Sequence 13505, A |
| 11 | 98.5 | 7.6 | 460 | 4 | US-09-489-039A-13505 | Sequence 13505, A |
| 12 | 97.5 | 7.5 | 355 | 4 | US-09-602-787A-530 | Sequence 530, App |
| 13 | 95 | 7.3 | 597 | 4 | US-09-252-991A-32657 | Sequence 32657, A |
| 14 | 94 | 7.2 | 483 | 4 | US-09-107-532A-4123 | Sequence 4123, Ap |
| 15 | 93.5 | 7.2 | 439 | 4 | US-09-252-991A-28709 | Sequence 28709, A |
| 16 | 93 | 7.1 | 327 | 4 | US-09-489-039A-8641 | Sequence 8641, Ap |
| 17 | 92.5 | 7.1 | 336 | 4 | US-09-252-991A-20404 | Sequence 20404, A |
| 18 | 92.5 | 7.1 | 452 | 3 | US-09-113-001C-3935 | Sequence 3935, Ap |
| 19 | 92.5 | 7.1 | 465 | 4 | US-09-198-452A-524 | Sequence 524, App |
| 20 | 92.5 | 7.1 | 465 | 4 | US-09-328-352-7056 | Sequence 7056, Ap |
| 21 | 92 | 7.1 | 664 | 4 | US-09-710-279-3292 | Sequence 3292, Ap |
| 22 | 92 | 7.1 | 394 | 4 | US-09-328-352-7056 | Sequence 7056, Ap |
| 23 | 91.5 | 7.0 | 476 | 3 | US-09-134-001C-3778 | Sequence 3778, Ap |
| 24 | 91 | 7.0 | 489 | 4 | US-09-328-352-5088 | Sequence 5088, Ap |
| 25 | 90.5 | 7.0 | 249 | 4 | US-09-282-218A-17 | Sequence 17, Appl |
| 26 | 90.5 | 7.0 | 373 | 4 | US-09-603-208A-174 | Sequence 174, App |
| 27 | 90 | 6.9 | 554 | 3 | US-09-134-001C-5109 | Sequence 5109, Ap |
| 28 | 89.5 | 6.9 | 293 | 4 | US-09-134-000C-6254 | Sequence 6254, Ap |
| 29 | 89 | 6.8 | 300 | 4 | US-09-393-634-17 | Sequence 17, Appl |
| 30 | 88.5 | 6.8 | 465 | 4 | US-09-438-185A-488 | Sequence 488, App |
| 31 | 88.5 | 6.8 | 719 | 4 | US-09-328-352-6274 | Sequence 6274, Ap |
| 32 | 88 | 6.8 | 452 | 4 | US-09-543-681A-6544 | Sequence 6544, Ap |
| 33 | 88 | 6.8 | 554 | 4 | US-09-540-236-2634 | Sequence 2634, Ap |
| 34 | 87.5 | 6.7 | 264 | 3 | US-09-134-001C-3780 | Sequence 3780, Ap |
| 35 | 87.5 | 6.7 | 355 | 4 | US-09-489-039A-9302 | Sequence 9302, Ap |
| 36 | 87.5 | 6.7 | 536 | 4 | US-09-328-352-4689 | Sequence 4689, Ap |
| 37 | 87.5 | 6.7 | 657 | 4 | US-09-252-991A-27682 | Sequence 27682, A |
| 38 | 87 | 6.7 | 455 | 4 | US-09-543-681A-7043 | Sequence 7043, Ap |
| 39 | 87 | 6.7 | 349 | 3 | US-09-489-039A-13185 | Sequence 13185, A |
| 40 | 86.5 | 6.6 | 349 | 4 | US-09-134-001C-4004 | Sequence 4004, Ap |
| 41 | 86.5 | 6.6 | 395 | 4 | US-09-328-352-5819 | Sequence 5819, Ap |
| 42 | 86.5 | 6.6 | 940 | 4 | US-09-328-352-8165 | Sequence 8165, Ap |
| 43 | 86 | 6.6 | 246 | 4 | US-09-710-279-900 | Sequence 900, App |
| 44 | 86 | 6.6 | 246 | 4 | US-09-710-279-1956 | Sequence 1956, App |
| 45 | 86 | 6.6 | 445 | 4 | US-09-602-777A-312 | Sequence 312, App |

ALIGNMENTS

```
RESULT 1
US-08-313-553-2
Sequence 2, Application US/08313553
Patent No. 5641650
;  GENERAL INFORMATION:
;  APPLICANT:  TURNER, George J.
;  APPLICANT:  BETLACH, Mary C.
;  TITLE OF INVENTION:  EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
;  TITLE OF INVENTION:  IN HALOBACTERIA
;  NUMBER OF SEQUENCES:  15
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  Walter H. Dreger
;  STREET:  4 Embarcadero Center, Suite 3400
;  CITY:  San Francisco
;  STATE:  California
;  COUNTRY:  USA
;  ZIP:  94111
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE:  Floppy disk
;  COMPUTER:  IBM PC compatible
;  OPERATING SYSTEM:  PC-DOS/MS-DOS
;  SOFTWARE:  PatentIn Release #1.0, Version #1.25
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER:  US/08/313,553
;  FILING DATE:
;  CLASSIFICATION:  435
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER:  US/08/038,662
;  FILING DATE:
;  ATTORNEY/AGENT INFORMATION:
;  NAME:  Dreger, Walter H.
;  REGISTRATION NUMBER:  24,190
;  REFERENCE/DOCKET NUMBER:  A-57669/WHD
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE:  (415) 781-1989
;  TELEFAX:  (415) 398-3249
;  INFORMATION FOR SEQ ID NO: 2:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH:  262 amino acids
;  TYPE:  amino acid
;  TOPOLOGY:  linear
;  MOLECULE TYPE:  protein
US-08-313-553-2

Query Match     13.6%;  Score 177;  DB 1;  Length 262;
Best Local Similarity   27.8%;  Pred. No. 3.4e-10;
Matches    69;  Conservative  40;  Mismatches 101;  Indels 38;  Gaps 13;

Qy   13 LPTFAAGGDLDASDYTGVSFW-----LVTAALLASTVFFFVERDRVS-AKWKTSLTVSGL 67
            |||          :|       |   | |   |:     |||   |:|  |:||   ||  |
```

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: October 25, 2005, 09:20:28

Title: US-09-847-513A-7
Perfect score: 1301
Sequence: 1 MKLLILGSVIALPTFAAGG............NKILFGLIWNVAVKESSNA 249

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1862994 seqs, 417510619 residues

Total number of hits satisfying chosen parameters: 1862994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*
```
 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
 4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
 5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
 6: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
 8: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
 9: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Search time 902 Seconds
(without alignments)
115.255 Million cell updates/sec

SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1301 | 100.0 | 249 | 10 | US-09-847-513A-7 | Sequence 7, Appl |
| 2 | 1301 | 100.0 | 249 | 17 | US-10-724-264A-3 | Sequence 3, Appli |
| 3 | 1296 | 99.6 | 250 | 10 | US-09-847-513A-5 | Sequence 5, Appli |
| 4 | 1296 | 99.6 | 250 | 10 | US-09-847-513A-45 | Sequence 45, Appl |
| 5 | 1296 | 99.6 | 250 | 17 | US-10-724-264A-37 | Sequence 37, Appl |
| 6 | 1296 | 99.5 | 251 | 17 | US-10-724-264A-165 | Sequence 165, App |
| 7 | 1294 | 99.4 | 249 | 17 | US-10-724-264A-167 | Sequence 167, App |
| 8 | 1293 | 99.4 | 249 | 17 | US-10-724-264A-101 | Sequence 101, App |
| 9 | 1292 | 99.3 | 249 | 17 | US-10-724-264A-163 | Sequence 163, App |
| 10 | 1292 | 99.3 | 250 | 17 | US-09-847-513A-11 | Sequence 11, Appl |
| 11 | 1291 | 99.2 | 250 | 10 | US-09-847-513A-33 | Sequence 33, Appl |
| 12 | 1290 | 99.2 | 250 | 17 | US-10-724-264A-111 | Sequence 111, App |
| 13 | 1289 | 99.1 | 250 | 17 | US-10-724-264A-65 | Sequence 65, Appl |
| 14 | 1289 | 99.1 | 250 | 17 | US-10-724-264A-69 | Sequence 69, App |
| 15 | 1289 | 99.1 | 250 | 17 | US-09-847-513A-77 | Sequence 77, Appl |
| 16 | 1289 | 99.1 | 250 | 17 | US-10-724-264A-83 | Sequence 83, Appl |
| 17 | 1289 | 99.1 | 250 | 17 | US-10-724-264A-85 | Sequence 85, Appl |
| 18 | 1289 | 99.1 | 250 | 17 | US-10-724-264A-87 | Sequence 87, Appl |
| 19 | 1288 | 99.1 | 250 | 10 | US-09-847-513A-47 | Sequence 47, Appl |
| 20 | 1288 | 99.0 | 251 | 17 | US-10-724-264A-39 | Sequence 39, Appl |
| 21 | 1287 | 98.9 | 250 | 17 | US-10-724-264A-81 | Sequence 81, Appl |
| 22 | 1286 | 98.8 | 250 | 17 | US-10-724-264A-61 | Sequence 61, Appl |
| 23 | 1286 | 98.8 | 250 | 17 | US-10-724-264A-73 | Sequence 73, Appl |
| 24 | 1286 | 98.8 | 250 | 17 | US-10-724-264A-91 | Sequence 91, Appl |
| 25 | 1285 | 98.8 | 250 | 17 | US-10-724-264A-59 | Sequence 59, Appl |
| 26 | 1285 | 98.8 | 250 | 17 | US-10-724-264A-89 | Sequence 89, Appl |
| 27 | 1284 | 98.7 | 250 | 10 | US-09-847-513A-35 | Sequence 35, Appl |
| 28 | 1284 | 98.6 | 250 | 17 | US-10-724-264A-27 | Sequence 27, Appl |
| 29 | 1283 | 98.6 | 250 | 17 | US-09-847-513A-39 | Sequence 39, Appl |
| 30 | 1283 | 98.6 | 250 | 17 | US-10-724-264A-71 | Sequence 71, Appl |
| 31 | 1283 | 98.6 | 251 | 17 | US-10-724-264A-97 | Sequence 97, Appl |
| 32 | 1283 | 98.6 | 251 | 17 | US-10-724-264A-31 | Sequence 31, Appl |
| 33 | 1282 | 98.5 | 250 | 17 | US-10-724-264A-63 | Sequence 63, Appl |
| 34 | 1282 | 98.5 | 250 | 17 | US-10-724-264A-75 | Sequence 75, Appl |
| 35 | 1282 | 98.5 | 250 | 17 | US-10-724-264A-95 | Sequence 95, Appl |
| 36 | 1281 | 98.5 | 249 | 17 | US-10-724-264A-25 | Sequence 25, Appl |
| 37 | 1280 | 98.4 | 250 | 17 | US-10-724-264A-79 | Sequence 79, Appl |
| 38 | 1280 | 98.4 | 250 | 17 | US-10-724-264A-93 | Sequence 93, Appl |
| 39 | 1279 | 98.3 | 250 | 17 | US-10-724-264A-109 | Sequence 109, App |
| 40 | 1279 | 98.2 | 250 | 17 | US-10-724-264A-103 | Sequence 103, App |
| 41 | 1277 | 98.2 | 250 | 17 | US-10-724-264A-113 | Sequence 113, App |
| 42 | 1275 | 98.0 | 250 | 17 | US-10-724-264A-67 | Sequence 67, Appl |
| 43 | 1274 | 97.9 | 250 | 10 | US-09-847-513A-41 | Sequence 41, Appl |
| 44 | 1274 | 97.9 | 251 | 17 | US-10-724-264A-33 | Sequence 33, Appl |
| 45 | 1273 | 97.8 | 249 | 17 | US-10-724-264A-105 | Sequence 105, App |

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
US-09-847-513A-7

; Sequence 7, Application US/09847513A
; Publication No. US2003010437 5A1
; GENERAL INFORMATION:
; APPLICANT: MBARI
; APPLICANT: Delong, Edward
; APPLICANT: Beja, Oded
; TITLE OF INVENTION: Light-driven energy generation using proteorhodopsin
; FILE REFERENCE: MBA-101
; CURRENT APPLICATION NUMBER: US/09/847,513A
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: 60/201,602
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 7
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Naturally occurring gamma proteobacterium
US-09-847-513A-7

Query Match 100.0%; Score 1301; DB 10; Length 249;
Best Local Similarity 100.0%; Pred. No. 5.2e-121;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 MKLLILGSVIALPTFAAGGGDLDASDYTGVSFWLVTAALLASTVFFVERDRVSAKWKT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MKLLILGSVIALPTFAAGGGDLDASDYTGVSFWLVTAALLASTVFFVERDRVSAKWKT 60

Qy   61 SLIVSGLVTGIAFWHYMYMRGVWIETGDSPTVFRYIDWLLTVPLLICEFYLILAAATNVA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 SLIVSGLVTGIAFWHYMYMRGVWIETGDSPTVFRYIDWLLTVPLLICEFYLILAAATNVA 120
```

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:          October 19, 2005, 08:08:47 ; Search time 3301 Seconds
                 (without alignments)
                 3655.059 Million cell updates/sec

Title:           US-09-847-513A-7
Perfect score:   1301
Sequence:        1 MKLLILGSVIALPTFAAGG..........NKILFGLLIWNVAVKESSNA 249

Scoring table:   BLOSUM62
                 Xgapop   10.0 ; Xgapext  0.5
                 Ygapop   10.0 ; Ygapext  0.5
                 Fgapop    6.0 ; Fgapext  7.0
                 Delop     6.0 ; Delext   7.0

Searched:        4708233 seqs, 2422607955 residues

Total number of hits satisfying chosen parameters:          9416466

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09847513/runat_19102005_090520_24129/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09847513 @CGN_1_1_5600 @runat_19102005_090520 24129 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10: gb_ro:*
                 11: gb_sts:*
                 12: gb_sy:*
                 13: gb_un:*
                 14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Query | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| c | | | | | | | |
| 1 | 1301 | 100.0 | 128758 | 1 | | AF279106 | Uncultu |
| 2 | 1296 | 99.6 | 751 | 1 | | AY601905 | Uncultu |
| 3 | 1296 | 99.6 | 753 | 1 | | AF349993 | Uncultu |
| 4 | 1293 | 99.4 | 751 | 1 | | AY210919 | Uncultu |
| 5 | 1290 | 99.2 | 751 | 1 | | AY250718 | Uncultu |
| 6 | 1289 | 99.1 | 750 | 1 | | AY210910 | Uncultu |
| 7 | 1289 | 99.1 | 751 | 1 | | AY210901 | Uncultu |
| 8 | 1289 | 99.1 | 751 | 1 | | AY210903 | Uncultu |
| 9 | 1289 | 99.1 | 751 | 1 | | AY210907 | Uncultu |
| 10 | 1289 | 99.1 | 751 | 1 | | AY210911 | Uncultu |
| 11 | 1289 | 99.1 | 751 | 1 | | AY210912 | Uncultu |
| 12 | 1288 | 99.0 | 753 | 1 | | AF349994 | Uncultu |
| 13 | 1287 | 98.9 | 753 | 1 | | AY210909 | Uncultu |
| 14 | 1286 | 98.8 | 751 | 1 | | AY210899 | Uncultu |
| 15 | 1286 | 98.8 | 751 | 1 | | AY210905 | Uncultu |
| 16 | 1286 | 98.8 | 751 | 1 | | AY210914 | Uncultu |
| 17 | 1285 | 98.8 | 751 | 1 | | AY210898 | Uncultu |
| 18 | 1285 | 98.8 | 751 | 1 | | AY210913 | Uncultu |
| 19 | 1284 | 98.7 | 753 | 1 | | AF349988 | Uncultu |
| 20 | 1283 | 98.6 | 751 | 1 | | AY210904 | Uncultu |
| 21 | 1283 | 98.6 | 751 | 1 | | AY210917 | Uncultu |
| 22 | 1283 | 98.6 | 753 | 1 | | AF349990 | Uncultu |
| 23 | 1282 | 98.5 | 751 | 1 | | AY210900 | Uncultu |
| 24 | 1282 | 98.5 | 751 | 1 | | AY210906 | Uncultu |
| 25 | 1282 | 98.5 | 751 | 1 | | AY210916 | Uncultu |
| 26 | 1281 | 98.5 | 748 | 1 | | AF349987 | Uncultu |
| 27 | 1280 | 98.4 | 750 | 1 | | AY210908 | Uncultu |
| 28 | 1280 | 98.4 | 751 | 1 | | AY210915 | Uncultu |
| 29 | 1279 | 98.3 | 751 | 1 | | AY250717 | Uncultu |
| 30 | 1277 | 98.2 | 751 | 1 | | AY250719 | Uncultu |
| 31 | 1275 | 98.0 | 751 | 1 | | AF349991 | Uncultu |
| 32 | 1274 | 97.9 | 753 | 1 | | AY210902 | Uncultu |
| 33 | 1273 | 97.8 | 748 | 1 | | AY250715 | Uncultu |
| 34 | 1273 | 97.8 | 751 | 1 | | AY598756 | Uncultu |
| 35 | 1273 | 97.8 | 751 | 1 | | AY250716 | Uncultu |
| 36 | 1273 | 97.8 | 757 | 1 | | AF349977 | Uncultu |
| 37 | 1272 | 97.8 | 751 | 1 | | AY210918 | Uncultu |
| 38 | 1270 | 97.6 | 753 | 1 | | AF349992 | Uncultu |
| 39 | 1268 | 97.5 | 751 | 1 | | AY250723 | Uncultu |
| 40 | 1267 | 97.4 | 747 | 1 | | AY250714 | Uncultu |
| 41 | 1267 | 97.4 | 753 | 1 | | AF349978 | Uncultu |
| 42 | 1267 | 97.4 | 753 | 1 | | AF349985 | Uncultu |
| 43 | 1262 | 97.0 | 757 | 1 | | AF349976 | Uncultu |
| 44 | 1260 | 96.8 | 753 | 1 | | AF349986 | Uncultu |
| 45 | 1259 | 96.8 | 751 | 1 | | AY250722 | Uncultu |

ALIGNMENTS

RESULT 1
AF279106/c
LOCUS       AF279106          128758 bp    DNA     linear   BCT 29-APR-2004
DEFINITION  Uncultured marine gamma proteobacterium EBAC31A08 BAC sequence.
ACCESSION   AF279106
VERSION     AF279106.2  GI:34112904
KEYWORDS    .
SOURCE      uncultured marine gamma proteobacterium EBAC31A08
  ORGANISM  uncultured marine gamma proteobacterium EBAC31A08
            Bacteria; Proteobacteria; Gammaproteobacteria; environmental
            samples.
REFERENCE   1  (bases 1 to 128758)
  AUTHORS   Beja,O., Aravind,L., Koonin,E.V., Suzuki,M.T., Hadd,A.,
            Nguyen,L.P., Jovanovich,S.B., Gates,C.M., Feldman,R.A.,
            Spudich,J.L., Spudich,E.N. and DeLong,E.F.
  TITLE     Bacterial rhodopsin: evidence for a new type of phototrophy in the
            sea
  JOURNAL   Science 289 (5486), 1902-1906 (2000)
  MEDLINE   20462260
   PUBMED   10988064
REFERENCE   2  (bases 1 to 128758)
  AUTHORS   Beja,O., Aravind,L., Koonin,E.V., Suzuki,M.T., Hadd,A.,
            Nguyen,L.P., Jovanovich,S.B., Gates,C.M., Feldman,R.A. and
            DeLong,E.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2000) R&D, Monterey Bay Aquarium Research
            Institute, 7700 Sandholdt Road, Moss Landing, CA 95039-0628, USA
REFERENCE   3  (bases 1 to 128758)

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: October 19, 2005, 08:05:27 ; Search time 452 Seconds
(without alignments)
3261.096 Million cell updates/sec

Title: US-09-847-513A-7
Perfect score: 1301
Sequence: 1 MKLLLLLGSVIALPTFAAGG........NKILFGLIIWNNVAVKESSNA 249

Searched: 4390206 seqs, 295987066 residues

Total number of hits satisfying chosen parameters: 8780412

Maximum DB seq length: 200000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Scoring table: BLOSUM62
Xgapop  10.0 ; Xgapext  0.5
Ygapop  10.0 ; Ygapext  0.5
Fgapop  6.0 ; Fgapext  7.0
Delop  6.0 ; Delext  7.0

Command line parameters: -DEV=xlp
-MODEL=frame+_p2n.model
-Q=/cgm2_1/USPTO_spool_p/US09847513/runat_19102005_090520_24121/app_query.fasta_1.391
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -HEAPSIZE=500 -MINLEN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000000
-USER=US09847513 @CGN_1_1_708 @runat_19102005_090520_24121 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: N Geneseq_16Dec04:*
2: geneseqm1980s:*
3: geneseqm1990s:*
4: geneseqm2001as:*
5: geneseqm2001bs:*
6: geneseqm2002as:*
7: geneseqm2002bs:*
8: geneseqm2003as:*
9: geneseqm2003bs:*
10: geneseqm2003cs:*
11: geneseqm2003ds:*
12: geneseqm2004as:*
13: geneseqm2004bs:*

## SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1301 | 100.0 | 747 | 6 | ABK24126 | Abk24126 DNA encod |
| 2 | 1301 | 100.0 | 750 | 13 | ABD82509 | Adg82509 Wild type |
| 3. | 1301 | 100.0 | 105184 | 6 | ABK24122 | ABK24122 Bacterial |
| 4 | 1296 | 99.6 | 750 | 6 | ABK24125 | Abk24125 DNA encod |
| 5 | 1296 | 99.6 | 750 | 6 | ABK24145 | Abk24145 DNA encod |
| 6 | 1296 | 99.6 | 753 | 13 | ADQ82543 | Adq82543 Wild type |
| 7 | 1294 | 99.5 | 750 | 13 | ADQ82671 | Adg82671 Mutant ma |
| 8 | 1293 | 99.4 | 750 | 13 | ABQ82673 | Adg82673 Mutant ma |
| 9 | 1293 | 99.3 | 751 | 13 | ADQ82607 | Adq82607 Wild type |
| 10 | 1292 | 99.3 | 750 | 6 | ABK24128 | Abk24128 DNA encod |
| 11 | 1292 | 99.3 | 750 | 13 | ADQ82669 | Adq82669 Wild type |
| 12 | 1291 | 99.2 | 750 | 6 | ABK24139 | ABK24139 DNA encod |
| 13 | 1290 | 99.2 | 751 | 13 | ADQ82617 | Adg82617 Wild type |
| 14 | 1289 | 99.1 | 750 | 13 | ADQ82589 | Adg82589 Wild type |
| 15 | 1289 | 99.1 | 751 | 13 | ADQ82607 | Adg82607 Wild type |
| 16 | 1289 | 99.1 | 751 | 13 | ADQ82591 | Adg82591 Wild type |
| 17 | 1289 | 99.1 | 751 | 13 | ADQ82575 | Adg82575 Wild type |
| 18 | 1289 | 99.1 | 751 | 13 | ADQ82583 | Adg82583 Wild type |
| 19 | 1289 | 99.1 | 751 | 13 | ADQ82593 | Adg82593 Wild type |
| 20 | 1288 | 99.0 | 750 | 6 | ADQ82571 | Adg82571 Wild type |
| 21 | 1288 | 99.0 | 753 | 13 | ADQ82545 | Adg82545 Wild type |
| 22 | 1287 | 98.9 | 750 | 13 | ADQ82587 | Adg82587 Wild type |
| 23 | 1286 | 98.8 | 751 | 13 | ADQ82597 | Adg82597 Wild type |
| 24 | 1286 | 98.8 | 751 | 13 | ADQ82567 | Adg82567 Wild type |
| 25 | 1286 | 98.8 | 751 | 13 | ADQ82579 | Adg82579 Wild type |
| 26 | 1285 | 98.8 | 751 | 13 | ADQ82595 | Adg82595 Wild type |
| 27 | 1285 | 98.8 | 751 | 13 | ADQ82565 | Adg82565 Wild type |
| 28 | 1284 | 98.7 | 750 | 6 | ABK24146 | Abk24146 DNA encod |
| 29 | 1284 | 98.7 | 753 | 13 | ADQ82533 | Adg82533 Wild type |
| 30 | 1283 | 98.6 | 750 | 6 | ABK24142 | ABK24142 DNA encod |
| 31 | 1283 | 98.6 | 751 | 13 | ADQ82577 | Adg82577 Wild type |
| 32 | 1283 | 98.6 | 751 | 13 | ADQ82603 | Adg82603 Wild type |
| 33 | 1283 | 98.6 | 753 | 13 | ADQ82537 | Adg82537 Wild type |
| 34 | 1282 | 98.5 | 751 | 13 | ADQ82601 | Adg82601 Wild type |
| 35 | 1282 | 98.5 | 751 | 13 | ADQ82669 | Adg82669 Wild type |
| 36 | 1282 | 98.5 | 751 | 13 | ADQ82581 | Adg82581 Wild type |
| 37 | 1281 | 98.5 | 748 | 13 | ADQ82531 | Adg82531 Wild type |
| 38 | 1280 | 98.4 | 750 | 13 | ADQ82585 | Adg82585 Wild type |
| 39 | 1280 | 98.4 | 751 | 13 | ADQ82599 | Adg82599 Wild type |
| 40 | 1279 | 98.3 | 751 | 13 | ADQ82615 | Adg82615 Wild type |
| 41 | 1277 | 98.2 | 751 | 13 | ADQ82609 | Adg82609 Wild type |
| 42 | 1277 | 98.2 | 751 | 13 | ADQ82619 | Adg82619 Wild type |
| 43 | 1275 | 98.1 | 751 | 13 | ADQ82573 | Adg82573 Wild type |
| 44 | 1274 | 97.9 | 750 | 6 | ABK24143 | Abk24143 DNA encod |
| 45 | 1274 | 97.9 | 753 | 13 | ADQ82539 | Adg82539 Wild type |

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## ALIGNMENTS

RESULT 1
ABK24126

ID   ABK24126 standard; DNA; 747 BP.
XX
AC   ABK24126;
XX
DT   07-AUG-2003 (revised)
DT   09-APR-2002 (first entry)
XX
DE   DNA encoding Proteorhodopsin from clone EBAC31A8 #2.
XX
KW   Proteorhodopsin; light-driven energy generator; targeted drug delivery;
KW   biocatalytic reactor; fuel cell; nano-machine; molecular switching;
KW   data storage; membrane potential; halophilic arachaea; ds.
XX
OS   Eubacteria.
XX
PN   WO200183701-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014394.
XX
PR   03-MAY-2000; 2000US-0201602P.
XX
PA   (MONT-) MONTEREY BAY AQUARIUM RES INST.
XX
PI   Delong EF, Beja O;
XX

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6

Run on:     October 19, 2005, 16:18:24 ; Search time 171 Seconds
(without alignments)
2382.649 Million cell updates/sec

Title:       US-09-847-513A-7
Perfect score: 1301
Sequence:   1 MKLLLIGSVIALPTFAAGG.........NKILFGLIIWNVAVKESSNA 249

Scoring table:  BLOSUM62
             Xgapop 10.0 , Xgapext 0.5
             Ygapop 10.0 , Ygapext 0.5
             Fgapop 6.0 , Fgapext 7.0
             Delop 6.0 , Delext 7.0

Searched:   1202784 seqs, 818138359 residues
             Maximum Match 100%
             Listing first 45 summaries
             Minimum Match 0%

Total number of hits satisfying chosen parameters:  2405568

Maximum DB seq length: 200000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
                Maximum Match 100%

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09847513//runat_19102005_090521_24155/app_query.fasta_1.391
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -THR_MIN=0 -ALIGN=15
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -TRANS=human40:cdi
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000000
-USER=US09847513 @CGN 1 1 105 @runat_19102005_090521_24155 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   Issued_Patents_NA:*
         1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
         2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
         3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
         4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
         5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
         6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 177 | 13.6 | 1254 | 1 | US-08-313-553-1 | Sequence 1, Appli |
| | 2 | 177 | 13.6 | 1254 | 3 | US-08-767-993-1 | Sequence 1, Appl |
| | 3 | 177 | 13.6 | 2147 | 3 | US-08-313-553-14 | Sequence 14, Appl |
| | 4 | 177 | 13.6 | 2147 | 3 | US-08-767-993-14 | Sequence 14, Appl |
| | 5 | 103.5 | 8.0 | 1830121 | 4 | US-09-557-884-1 | Sequence 1, Appli |
| | 6 | 103.5 | 8.0 | 1830121 | 4 | US-09-643-990A-1 | Sequence 1, Appli |
| | 7 | 102.5 | 7.9 | 1401 | 4 | US-09-252-991A-10908 | Sequence 10908, A |
| | 8 | 102.5 | 7.9 | 1440 | 4 | US-09-252-991A-11041 | Sequence 11041, A |
| | 9 | 102 | 7.8 | 1338 | 4 | US-09-603-208A-261 | Sequence 261, App |
| c | 10 | 99 | 7.6 | 501 | 4 | US-09-252-991A-11149 | Sequence 11149, A |
| c | 11 | 99 | 7.6 | 1368 | 4 | US-09-602-787A-529 | Sequence 529, App |
| | 12 | 98.5 | 7.6 | 1383 | 4 | US-09-489-039A-6334 | Sequence 6334, Ap |
| | 13 | 97.5 | 7.5 | 1068 | 4 | US-09-248-796A-155 | Sequence 155, App |
| | 14 | 97.5 | | 42325 | 4 | US-08-311-731A-131 | Sequence 131, App |
| | 15 | 96 | 7.4 | 4403765 | 3 | US-09-103-840A-2 | Sequence 2, Appli |
| | 16 | 96 | 7.4 | 4411529 | 3 | US-09-103-840A-1 | Sequence 1, Appli |
| | 17 | 95.5 | 7.4 | 1470 | 4 | US-09-328-352-962 | Sequence 962, App |
| | 18 | 95 | 7.3 | 1259 | 2 | US-08-672-814D-12 | Sequence 12, Appl |
| | 19 | 95 | 7.3 | 1259 | 2 | US-09-333-696-12 | Sequence 12, Appl |
| | 20 | 95 | 7.3 | 1794 | 4 | US-09-252-991A-16086 | Sequence 16086, A |
| | 21 | 95 | 7.3 | 1878 | 4 | US-09-252-991A-16516 | Sequence 16516, A |
| c | 22 | 94 | 7.2 | 1452 | 4 | US-09-107-532A-469 | Sequence 469, App |
| | 23 | 93.5 | 7.2 | 1320 | 4 | US-09-252-991A-12138 | Sequence 12138, A |
| c | 24 | 93.5 | 7.2 | 8536 | 4 | US-08-956-171E-278 | Sequence 278, App |
| c | 25 | 93.5 | 7.2 | 8536 | 4 | US-09-781-986A-278 | Sequence 278, App |
| | 26 | 93 | 7.2 | 984 | 4 | US-09-489-039A-15987 | Sequence 15987, A |
| | 27 | 93 | 7.1 | 1776 | 4 | US-09-252-991A-1470 | Sequence 1470, Ap |
| | 28 | 93 | 7.1 | 3164 | 1 | US-08-188-228-43 | Sequence 43, Appl |
| | 29 | 93 | 7.1 | 3164 | 1 | US-08-332-643-43 | Sequence 43, Appl |
| | 30 | 93 | 7.1 | 3164 | 1 | US-08-332-638-49 | Sequence 49, Appl |
| | 31 | 93 | 7.1 | 4000 | 4 | US-09-016-434-1434 | Sequence 1434, Ap |
| | 32 | 93 | 7.1 | 4000 | 4 | US-09-023-655-1452 | Sequence 1452, Ap |
| c | 33 | 93 | 7.1 | 23439 | 4 | US-08-956-171E-38 | Sequence 38, Appl |
| c | 34 | 93 | 7.1 | 23439 | 4 | US-09-781-986A-38 | Sequence 38, Appl |
| | 35 | 92.5 | 7.1 | 1011 | 4 | US-09-252-991A-3833 | Sequence 3833, Ap |
| | 36 | 92.5 | 7.1 | 1290 | 4 | US-09-252-991A-3792 | Sequence 3792, Ap |
| | 37 | 92.5 | 7.1 | 1359 | 3 | US-09-134-001C-1098 | Sequence 1098, Ap |
| | 38 | 92.5 | 7.1 | 1995 | 4 | US-09-328-352-2930 | Sequence 2930, Ap |
| | 39 | 92.5 | 7.1 | 2156 | 4 | US-09-710-279-4428 | Sequence 4428, Ap |
| c | 40 | 92.5 | 7.1 | 2742 | 4 | US-09-252-991A-3678 | Sequence 3678, Ap |
| c | 41 | 92.5 | 7.1 | 1230025 | 4 | US-09-198-452A-1 | Sequence 1, Appli |
| c | 42 | 92.5 | 7.1 | 1830121 | 4 | US-09-557-884-1 | Sequence 1, Appli |
| c | 43 | 92.5 | 7.1 | 1830121 | 4 | US-09-643-990A-1 | Sequence 1, Appli |
| c | 44 | 92 | 7.1 | 1180 | 3 | US-09-710-279-3291 | Sequence 3291, Ap |
| c | 45 | 92 | 7.1 | 1428 | 3 | US-09-134-001C-941 | Sequence 941, App |

## ALIGNMENTS

RESULT 1
US-08-313-553-1
Sequence 1, Application US/08313553
Patent No. 5641650

GENERAL INFORMATION:

APPLICANT: TURNER, George J.
APPLICANT: BETLACH, Mary C.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
TITLE OF INVENTION: IN HALOBACTERIA
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Walter H. Dreger
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,553
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,553
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/038,662
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-57669/WHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: October 19, 2005, 17:38:44 ; Search time 693 Seconds
(without alignments)
2519.111 Million cell updates/sec

Title: US-09-847-513A-7
Perfect score: 1301
Sequence: 1 MKLLILGSVIALPTFAAGG..........NKILFGLIIWNVAVKESSNA 249

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 8766186 seqs, 3505510206 residues

Total number of hits satisfying chosen parameters: 17532372

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09847513/runat_19102005_090522_24184/app_query.fasta_1.391
-DS=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09847513 @CGN_1_1_723 @runat_19102005_090522_24184
-NCPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
3: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US0A_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US0B_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US0C_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US0D_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US0E_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US0F_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US0G_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

---

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1301 | 100.0 | 747 | 10 | US-10-724-264A-4 | Sequence 6, Appli |
| 2 | 1301 | 100.0 | 751 | 21 | US-10-724-264A-4 | Sequence 4, Appli |
| 3 | 1301 | 100.0 | 105184 | | | Sequence 1, Appli |
| 4 | 1296 | 99.6 | 750 | 10 | US-09-847-513A-7 | Sequence 44, Appli |
| 5 | 1296 | 99.6 | 750 | 10 | US-09-847-513A-44 | Sequence 44, Appli |
| 6 | 1296 | 99.6 | 753 | 21 | US-09-847-513A-38 | Sequence 38, Appli |
| 7 | 1294 | 99.5 | 750 | 21 | US-10-724-264A-166 | Sequence 166, App |
| 8 | 1293 | 99.4 | 750 | 21 | US-10-724-264A-168 | Sequence 168, Appl |
| 9 | 1293 | 99.4 | 751 | 21 | US-10-724-264A-102 | Sequence 102, App |
| 10 | 1292 | 99.3 | 750 | 10 | US-09-847-513A-10 | Sequence 10, Appl |
| 11 | 1292 | 99.3 | 750 | 21 | US-10-724-264A-164 | Sequence 164, Appl |
| 12 | 1291 | 99.2 | 750 | 10 | US-09-847-513A-32 | Sequence 32, Appl |
| 13 | 1291 | 99.2 | 751 | 21 | US-10-724-264A-112 | Sequence 112, Appl |
| 14 | 1289 | 99.1 | 750 | 21 | US-10-724-264A-84 | Sequence 84, Appl |
| 15 | 1289 | 99.1 | 751 | 21 | US-10-724-264A-66 | Sequence 66, Appl |
| 16 | 1289 | 99.1 | 751 | 21 | US-10-724-264A-70 | Sequence 70, Appl |
| 17 | 1289 | 99.1 | 751 | 21 | US-10-724-264A-78 | Sequence 78, Appl |
| 18 | 1289 | 99.1 | 751 | 21 | US-10-724-264A-86 | Sequence 86, Appl |
| 19 | 1289 | 99.1 | 751 | 21 | US-10-724-264A-88 | Sequence 88, Appl |
| 20 | 1288 | 99.0 | 750 | 10 | US-09-847-513A-46 | Sequence 46, Appl |
| 21 | 1288 | 99.0 | 753 | 21 | US-10-724-264A-40 | Sequence 40, Appl |
| 22 | 1287 | 98.9 | 751 | 21 | US-10-724-264A-82 | Sequence 82, Appl |
| 23 | 1286 | 98.8 | 751 | 21 | US-10-724-264A-62 | Sequence 62, Appl |
| 24 | 1286 | 98.8 | 751 | 21 | US-10-724-264A-74 | Sequence 74, Appl |
| 25 | 1286 | 98.8 | 751 | 21 | US-10-724-264A-92 | Sequence 92, Appl |
| 26 | 1285 | 98.8 | 751 | 21 | US-10-724-264A-60 | Sequence 60, Appl |
| 27 | 1285 | 98.8 | 751 | 21 | US-10-724-264A-90 | Sequence 90, Appl |
| 28 | 1284 | 98.7 | 750 | 10 | US-09-847-513A-34 | Sequence 34, Appl |
| 29 | 1284 | 98.7 | 753 | 21 | US-10-724-264A-28 | Sequence 28, Appl |
| 30 | 1283 | 98.6 | 750 | 10 | US-09-847-513A-72 | Sequence 72, Appl |
| 31 | 1283 | 98.6 | 751 | 21 | US-10-724-264A-72 | Sequence 72, Appl |
| 32 | 1283 | 98.6 | 751 | 21 | US-10-724-264A-98 | Sequence 98, Appl |
| 33 | 1283 | 98.6 | 753 | 21 | US-10-724-264A-32 | Sequence 32, Appl |
| 34 | 1282 | 98.5 | 751 | 21 | US-10-724-264A-64 | Sequence 64, Appl |
| 35 | 1282 | 98.5 | 751 | 21 | US-10-724-264A-76 | Sequence 76, Appl |
| 36 | 1282 | 98.5 | 751 | 21 | US-10-724-264A-96 | Sequence 96, Appl |
| 37 | 1281 | 98.5 | 748 | 21 | US-10-724-264A-26 | Sequence 26, Appl |
| 38 | 1280 | 98.4 | 750 | 21 | US-10-724-264A-80 | Sequence 80, Appl |
| 39 | 1280 | 98.4 | 751 | 21 | US-10-724-264A-94 | Sequence 94, Appl |
| 40 | 1279 | 98.3 | 751 | 21 | US-10-724-264A-110 | Sequence 110, App |
| 41 | 1277 | 98.2 | 751 | 21 | US-10-724-264A-104 | Sequence 104, App |
| 42 | 1277 | 98.2 | 751 | 21 | US-10-724-264A-114 | Sequence 114, App |
| 43 | 1275 | 98.0 | 751 | 21 | US-10-724-264A-68 | Sequence 68, Appl |
| 44 | 1274 | 97.9 | 750 | 10 | US-09-847-513A-40 | Sequence 40, Appl |
| 45 | 1274 | 97.9 | 753 | 21 | US-10-724-264A-34 | Sequence 34, Appl |

ALIGNMENTS

RESULT 1
US-09-847-513A-6

; Sequence 6, Application US/09847513A
; Publication No. US20030104375A1
; GENERAL INFORMATION:
; APPLICANT: MBARI
; APPLICANT: DeLong, Edward
; APPLICANT: Beja, Oded
; TITLE OF INVENTION: Light-driven energy generation using proteorhodopsin
; FILE REFERENCE: MBA-101
; CURRENT APPLICATION NUMBER: US/09/847,513A
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: 60/201,602
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0

us-09-847-513a-7.p2n.rst

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: October 19, 2005, 13:47:19 ; Search time 3084 Seconds
(without alignments)
3073.282 Million cell updates/sec

Title:
Perfect score: 1301
Sequence: 1 MKLILILGSVIALPTFAAGG.........NKILFGLIIWNVAVKESSNA 249

Scoring table:

| | Xgapext | |
|---|---|---|
| BLOSUM62 | | |
| Xgapop 10.0 , | Xgapext | 0.5 |
| Ygapop 10.0 , | Ygapext | 0.5 |
| Fgapop 6.0 , | Fgapext | 7.0 |
| Delop 6.0 , | Delext | 7.0 |

Searched: 34235544 seqs, 1903134700 residues
Maximum Match 100%
Listing first 45 summaries

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

US-09-847-513A-7

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/Cgn2_1/USPTO_spool_p/US09847513/runat_19102005_090521_24141/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=pct -MINLEN=0 -MAXLEN=200000000
-USER=US09847513 @CGN 1 1 5180 @runat_19102005_090521_24141 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -YGAPOP=10 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| c 1 | 173.5 | 13.3 | 809 | 2 | AW342219 | AW342219 GthEST105 |
| 2 | 164 | 12.6 | 627 | 1 | AJ639139 | AJ639139 AJ639139 |
| 3 | 155 | 11.9 | 752 | 1 | AJ639344 | AJ639344 AJ639344 |
| 4 | 148.5 | 11.4 | 715 | 7 | CF259014 | CF259014 j321r.abl |
| 5 | 142 | 10.9 | 661 | 1 | AJ637035 | AJ637035 AJ637035 |
| 5 | 142 | 10.9 | 702 | 1 | AJ639451 | AJ639451 AJ639451 |
| 7, | 140 | 10.8 | 631 | 1 | AJ639452 | AJ639452 AJ639452 |
| 8 | 140 | 10.8 | 668 | 1 | AJ638560 | AJ638560 AJ638560 |
| 9 | 138 | 10.6 | 660 | 8 | BZ891866 | BZ891866 Hm10_0135 |

| 10 | 132.5 | 10.2 | 607 | 1 | AJ638781 | AJ638781 AJ638781 |
| 11 | 130 | 10.0 | 872 | 6 | CA764330 | CA764330 AF53-Rpf |
| 12 | 128.5 | 9.9 | 931 | 7 | CO008926 | CO008926 EST797261 |
| 13 | 126.5 | 9.7 | 780 | 7 | CO012533 | CO012533 EST800868 |
| 14 | 126.5 | 9.7 | 871 | 7 | CO011377 | CO011377 EST799712 |
| 15 | 126.5 | 9.7 | 879 | 7 | CO014037 | CO014037 EST802372 |
| 16 | 126.5 | 9.7 | 883 | 7 | CO025968 | CO025968 EST804352 |
| 17 | 126.5 | 9.7 | 912 | 7 | CO004549 | CO004549 EST792884 |
| 18 | 125.5 | 9.6 | 609 | 1 | AJ638428 | AJ638428 AJ638428 |
| 19 | 121 | 9.3 | 720 | 7 | CO012879 | CO012879 EST801214 |
| 20 | 121 | 9.3 | 828 | 7 | CO034658 | CO034658 EST813042 |
| 21 | 121 | 9.3 | 875 | 7 | CO047790 | CO047790 EST793125 |
| 22 | 119.5 | 9.2 | 717 | 7 | CK448874 | CK448874 negBpl.1G |
| 23 | 119 | 9.1 | 713 | 7 | CK446419 | CK446419 pncs909aH |
| 24 | 116 | 8.9 | 1662 | 9 | CG756213 | CG756213 P051-3-B1 |
| 25 | 115.5 | 8.9 | 603 | 1 | AJ636093 | AJ636093 AJ636093 |
| 26 | 115.5 | 8.9 | 672 | 1 | AJ638198 | AJ638198 AJ638198 |
| 27 | 115.5 | 8.9 | 687 | 1 | AJ636129 | AJ636129 AJ636129 |
| 28 | 115.5 | 8.9 | 716 | 7 | CK446337 | CK446337 pncs910aH |
| 29 | 115 | 8.8 | 701 | 7 | CV111522 | CV111522 L2P08E02 |
| 30 | 114.5 | 8.8 | 592 | 7 | CK448876 | CK448876 negspl2C7 |
| 31 | 114 | 8.8 | 582 | 7 | CF643569 | CF643569 K10_B09 F |
| 32 | 112.5 | 8.7 | 703 | 2 | BF251110 | BF251110 EST418368 |
| 33 | 112.5 | 8.7 | 715 | 7 | CO012144 | CO012144 EST800479 |
| 34 | 112 | 8.6 | 612 | 7 | CF639423 | CF639423 D15_E03 F |
| c 35 | 111.5 | 8.6 | 941 | 2 | BF683746 | BF683746 602140026 |
| 36 | 111.5 | 8.6 | 952 | 5 | BQ890217 | BQ890217 AGENCOURT |
| 37 | 110.5 | 8.5 | 1507 | 9 | CG756860 | CG756860 P051-4-H0 |
| 38 | 108.5 | 8.3 | 585 | 5 | BZ305206 | BZ305206 KD3382.q1 |
| 39 | 108 | 8.3 | 413 | 7 | CF642219 | CF642219 D49_A08 F |
| 40 | 108 | 8.3 | 882 | 7 | CF699316 | CF699316 CCACH82TR |
| 41 | 107.5 | 8.3 | 888 | 6 | CK448876 | CK448876 negsp12C7 |
| c 42 | 107.5 | 8.3 | 947 | 6 | CD458575 | CD458575 Fg08_11d0 |
| 43 | 107.5 | 8.3 | 969 | 6 | AG129346 | AG129346 Pan_trogl |
| c 44 | 104 | 8.0 | 565 | 6 | CD489118 | CD489118 T18_H07 F |
| 45 | 104 | 8.0 | 646 | 7 | CF641333 | CF641333 D38_E03 F |

ALIGNMENTS

RESULT 1
AW342219/c
LOCUS    AW342219      809 bp   mRNA   linear   EST 01-JAN-2001
DEFINITION  GthEST105 Guillardia theta Lambda Zap II cDNA Library Guillardia
          theta cDNA clone Est272_7D, mRNA sequence.
ACCESSION   AW342219
VERSION     AW342219.1   GI:12000594
KEYWORDS    EST.
SOURCE      Guillardia theta
  ORGANISM  Guillardia theta
          Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
REFERENCE   1 (bases 1 to 809)
  AUTHORS   Fraunholz,M., Duebel,J., Wastl,J., Zauner,S. and Maier,U.-G.
  TITLE     EST Database of the cryptomonad alga: Guillardia theta
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Maier, U.-G.
          Department of Cell Biology and Applied Botany
          Philipps-University Marburg
          Karl-von-Frisch-Strasse, D-35043 Marburg, Germany
          Tel: ++49 6421 282 2057
          Fax: ++49 6421 282 1543
          Email: maier@mailer.uni-marburg.de.
FEATURES          Location/Qualifiers
    source       1..809
            /organism="Guillardia theta"
            /mol_type="mRNA"
            /strain="CCMP327"
            /db_xref="taxon:55529"
            /clone="Est272_7D"
            /clone_lib="Guillardia theta Lambda Zap II cDNA Library"
            /lab_host="SOLR"
            /note="Vector: Lambda ZAP II; Site_1: EcoRI"
ORIGIN

5

OM nucleic - nucleic search, using sw model

Run on: February  9, 2004, 14:04:19 ; Search time 3035 Seconds
      (without alignments)
      10069.025 Million cell updates//sec

GenCore version 5.1.6
Copyright (c) 1993 - 2004  Compugen Ltd.

Title: US-09-847-513A-6
Perfect score: 747
Sequence: 1 atgaaattactactgatatt.........ttaaagaatcctctaatgct 747

Scoring table: IDENTITY_NUC
  Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
  Maximum Match 100%
  Listing first 45 summaries

Database :

  GenEmbl:*
 1: gb_ba:*
 2: gb_htg:*
 3: gb_in:*
 4: gb_om:*
 5: gb_ov:*
 6: gb_pat:*
 7: gb_ph:*
 8: gb_pl:*
 9: gb_pr:*
 10: gb_ro:*
 11: gb_sts:*
 12: gb_sy:*
 13: gb_un:*
 14: gb_vi:*
 15: gb_vr:*
 16: em_ba:*
 17: em_fun:*
 18: em_hum:*
 19: em_in:*
 20: em_mu:*
 21: em_om:*
 22: em_or:*
 23: em_ov:*
 24: em_pat:*
 25: em_ph:*
 26: em_pl:*
 27: em_ro:*
 28: em_sts:*
 29: em_un:*
 30: em_vi:*
 31: em_htg_hum:*
 32: em_htg_inv:*
 33: em_htg_other:*
 34: em_htg_mus:*
 35: em_htg_pln:*
 36: em_htg_rod:*
 37: em_htg_mam:*
 38: em_htg_vrt:*
 39: em_hgo_hum:*
 40: em_hgo_mus:*
 41: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| c | | | | | | |
| 1 | 747 | 100.0 | 105184 | 1 | AF279106 | Unculture |
| 2 | 740.8 | 99.2 | 750 | 1 | AY210909 | Uncultre |
| 3 | 739.2 | 99.0 | 751 | 1 | AY210913 | Uncultre |
| 4 | 739.2 | 99.0 | 751 | 1 | AY210919 | Uncultre |
| 5 | 737.6 | 98.7 | 751 | 1 | AY210910 | Uncultre |
| 6 | 736 | 98.5 | 750 | 1 | AY210908 | Uncultre |
| 7 | 734.4 | 98.3 | 751 | 1 | AY210918 | Uncultre |
| 8 | 734.4 | 98.3 | 751 | 1 | AY250714 | Uncultre |
| 9 | 734.4 | 98.3 | 751 | 1 | AY250717 | Uncultre |
| 10 | 734.2 | 98.3 | 751 | 1 | AF349987 | Uncultre |
| 11 | 731.4 | 97.9 | 748 | 1 | AY250715 | Uncultre |
| 12 | 731.2 | 97.9 | 751 | 1 | AY250716 | Uncultre |
| 13 | 728 | 97.5 | 753 | 1 | AF349990 | Uncultre |
| 14 | 728 | 97.5 | 753 | 1 | AY210993 | Uncultre |
| 15 | 724.8 | 97.1 | 751 | 1 | AY210917 | Uncultre |
| 16 | 713.6 | 95.5 | 753 | 1 | AF349988 | Uncultre |
| 17 | 705.6 | 94.5 | 753 | 1 | AF349986 | Uncultre |
| 18 | 705.6 | 94.5 | 753 | 1 | AF349992 | Uncultre |
| 19 | 704 | 94.2 | 753 | 1 | AF349985 | Uncultre |
| 20 | 700.8 | 93.8 | 753 | 1 | AF349991 | Uncultre |
| 21 | 697.6 | 93.4 | 757 | 1 | AF349977 | Uncultre |
| 22 | 696 | 93.2 | 753 | 1 | AF349994 | Uncultre |
| 23 | 694.4 | 93.0 | 753 | 1 | AF349978 | Uncultre |
| 24 | 694.4 | 93.0 | 751 | 1 | AF349976 | Uncultre |
| 25 | 689.6 | 92.3 | 753 | 1 | AF349983 | Uncultre |
| 26 | 689.4 | 92.3 | 748 | 1 | AF349989 | Uncultre |
| 27 | 688 | 92.1 | 753 | 1 | AF349984 | Uncultre |
| 28 | 667.2 | 89.3 | 751 | 1 | AY210911 | Uncultre |
| 29 | 659.2 | 88.2 | 751 | 1 | AY210905 | Uncultre |
| 30 | 659.2 | 88.2 | 751 | 1 | AY210914 | Uncultre |
| 31 | 657.6 | 88.0 | 751 | 1 | AY250718 | Uncultre |
| 32 | 654.4 | 87.6 | 751 | 1 | AY250719 | Uncultre |
| 33 | 652.8 | 87.4 | 751 | 1 | AY210915 | Uncultre |
| 34 | 651.2 | 87.2 | 753 | 1 | AY250720 | Uncultre |
| 35 | 648 | 86.7 | 751 | 1 | AY210901 | Uncultre |
| 36 | 648 | 86.7 | 751 | 1 | AY210903 | Uncultre |
| 37 | 648 | 86.7 | 751 | 1 | AY210904 | Uncultre |
| 38 | 648 | 86.7 | 751 | 1 | AY210906 | Uncultre |
| 39 | 648 | 86.7 | 751 | 1 | AY210907 | Uncultre |
| 40 | 646.4 | 86.5 | 751 | 1 | AY210898 | Uncultre |
| 41 | 646.4 | 86.5 | 751 | 1 | AY210900 | Uncultre |
| 42 | 644.8 | 86.3 | 751 | 1 | AY210899 | Uncultre |
| 43 | 644.8 | 86.3 | 751 | 1 | AY210902 | Uncultre |
| 44 | 644.8 | 86.3 | 751 | 1 | AY210912 | Uncultre |
| 45 | 644.8 | 86.3 | 751 | 1 | AY210916 | Uncultre |

ALIGNMENTS

RESULT 1
AF279106/c
LOCUS  AF279106  105184 bp  DNA  linear  BCT 23-OCT-2000
DEFINITION Uncultured proteobacterium EBAC31A08 clone BAC EBAC31A08, complete
    sequence.
ACCESSION AF279106
VERSION  AF279106.1  GI:9971876
KEYWORDS .
SOURCE  uncultured proteobacterium EBAC31A08
 ORGANISM uncultured proteobacterium EBAC31A08
    Bacteria; Proteobacteria; environmental samples.
REFERENCE 1  (bases 1 to 105184)
 AUTHORS Beja,O., Aravind,L., Koonin,E.V., Suzuki,M.T., Hadd,A.,
    Nguyen,L.P., Jovanovich,S.B., Gates,C.M., Feldman,R.A.,
    Spudich,J.L., Spudich,B.N. and DeLong,E.F.

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:          February  9, 2004, 14:04:18 ; Search time 273 Seconds
                 (without alignments)
                 7386.374 Million cell updates/sec

Title:           US-09-847-513A-6.
Perfect score:   747
Sequence:        1 atgaaattattactgatatt.........ttaaagaatcttccaaatgct 747

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 ; Gapext 1.0

Searched:        2552756 seqs, 134971 9017 residues

Total number of hits satisfying chosen parameters:    5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_19Jun03.*
    1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
    2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
    3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
    4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
    5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
    6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
    7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
    8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
    9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
   10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
   11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
   12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
   13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
   14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
   15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
   16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
   17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
   18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
   19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
   20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
   21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
   22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
   23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
   24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
   25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 747 | 100.0 | 747 | 24 | ABK24126 | DNA encoding Prote |
| c  2 | 747 | 100.0 | 105184 | 24 | ABK24122 | Bacterial artifici |
| 3 | 744 | 99.6 | 750 | 24 | ABK24125 | DNA encoding Prote |
| 4 | 739.2 | 99.0 | 750 | 24 | ABK24139 | DNA encoding Prote |
| 5 | 737.6 | 98.7 | 750 | 24 | ABK24128 | DNA encoding Prote |
| 6 | 728 | 97.5 | 750 | 24 | ABK24142 | DNA encoding Prote |
| 7 | 728 | 97.5 | 750 | 24 | ABK24145 | DNA encoding Prote |
| 8 | 713.6 | 95.5 | 750 | 24 | ABK24140 | DNA encoding Prote |
| 9 | 705.6 | 94.5 | 750 | 24 | ABK24138 | DNA encoding Prote |
| 10 | 705.6 | 94.5 | 750 | 24 | ABK24144 | DNA encoding Prote |
| 11 | 704 | 94.2 | 750 | 24 | ABK24137 | DNA encoding Prote |
| 12 | 700.8 | 93.8 | 750 | 24 | ABK24143 | DNA encoding Prote |
| 13 | 697.6 | 93.4 | 750 | 24 | ABK24129 | DNA encoding Prote |
| 14 | 696 | 93.2 | 750 | 24 | ABK24146 | DNA encoding Prote |
| 15 | 694.4 | 93.0 | 750 | 24 | ABK24127 | DNA encoding Prote |
| 16 | 694.4 | 93.0 | 750 | 24 | ABK24130 | DNA encoding Prote |
| 17 | 689.6 | 92.3 | 750 | 24 | ABK24141 | DNA encoding Prote |
| 18 | 688 | 92.1 | 750 | 24 | ABK24135 | DNA encoding Prote |
| 19 | 688 | 92.1 | 750 | 24 | ABK24136 | DNA encoding Prote |
| 20 | 459 | 61.4 | 753 | 24 | ABK24132 | DNA encoding Prote |
| 21 | 459 | 61.4 | 753 | 24 | ABK24154 | DNA encoding Prote |
| 22 | 457.4 | 61.2 | 753 | 24 | ABK24147 | DNA encoding Prote |
| 23 | 457.4 | 61.2 | 753 | 24 | ABK24151 | DNA encoding Prote |
| 24 | 455.8 | 61.0 | 753 | 24 | ABK24134 | DNA encoding Prote |
| 25 | 455.8 | 61.0 | 753 | 24 | ABK24150 | DNA encoding Prote |
| 26 | 454.2 | 60.8 | 753 | 24 | ABK24149 | DNA encoding Prote |
| 27 | 452.6 | 60.6 | 753 | 24 | ABK24133 | DNA encoding Prote |
| 28 | 452.6 | 60.6 | 753 | 24 | ABK24148 | DNA encoding Prote |
| 29 | 452.6 | 60.6 | 753 | 24 | ABK24152 | DNA encoding Prote |
| 30 | 452.6 | 60.6 | 753 | 24 | ABK24155 | DNA encoding Prote |
| 31 | 451 | 60.4 | 753 | 24 | ABK24153 | DNA encoding Prote |
| 32 | 451 | 60.2 | 753 | 24 | ABK24131 | DNA encoding Prote |
| 33 | 449.4 | 60.2 | 753 | 24 | AAS46608 | Tumour suppressor |
| 34 | 448.6 | 6.5 | 6767 | 22 | ABL70427 | Chemically treated |
| 35 | 448.6 | 6.5 | 7131 | 24 | AAS61360 | Human gene, regulat |
| 36 | 448.6 | 6.5 | 7131 | 24 | ABK31450 | Signal transductio |
| 37 | 447.4 | 6.3 | 7131 | 24 | ABL33103 | Human immune syste |
| 38 | 45 | 6.0 | 7758 | 24 | ABL32605 | Human immune syste |
| 39 | 45 | 6.0 | 11155 | 24 | ABL33943 | Human immune syste |
| 40 | 45 | 6.0 | 12138 | 24 | ABK28336 | Human immune syste |
| 41 | 44.8 | 6.0 | 12138 | 24 | ABL34074 | DNA transcription |
| 42 | 44.8 | 6.0 | 8297 | 24 | ABL34074 | Human immune syste |
| 43 | 44.6 | 6.0 | 56153 | 22 | AAS46794 | Tumour suppressor |
| 44 | 44.6 | 6.0 | 390 | 24 | ABA97217 | Ribosomal protein |
| 45 | 44.6 | 6.0 | 9652 | 24 | ABL32909 | Human immune syste |

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
ABK24126

ID   ABK24126 standard; DNA; 747 BP.
XX
AC   ABK24126;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding Proteorhodopsin from clone EBAC31A8 #2.
XX
KW   Proteorhodopsin; light-driven energy generator;
KW   targeted drug delivery; biocatalytic reactor; fuel cell;
KW   nano-machine; molecular switching; data storage; membrane potential;
KW   halophilic archaea; ds.
XX
OS   Bacteriaceae.
XX
PN   WO200183701-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US14394.
XX
PR   03-MAY-2000; 2000US-201602P.
XX
PA   (MONT-) MONTEREY BAY AQUARIUM RES INST.
XX
PI   Delong EF,  Beja O;
XX
DR   WPI; 2002-114151/15.
DR   P-PSDB; AAU81187.

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         February 9, 2004, 14:04:19 ; Search time 67 Seconds
                                             (without alignments)
                                             4921.093 Million cell updates/sec

Title:          US-09-847-513A-6
Perfect score:  747
Sequence:       1 atgaaattattactgatatt.........ttaaagaatcttctaatgct 747
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters:   1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued_Patents_NA:*
         1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
         2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
         3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
         4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
         5: /cgn2_6/ptodata/1/ina/pCTUS_COMB.seq:*
         6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 43.2 | 5.8 | 1497 | 4 | US-09-220-132-94 | Sequence 94, Appli |
| 2 | 42.2 | 5.6 | 6794 | 4 | US-09-491-356C-2 | Sequence 2, Appli |
| 3 | 42 | 5.6 | 9793 | 1 | US-08-470-202-56 | Sequence 56, Appl |
| 4 | 42 | 5.6 | 9793 | 2 | US-08-471-770-56 | Sequence 56, Appl |
| 5 | 42 | 5.6 | 9793 | 2 | US-08-468-059-56 | Sequence 56, Appl |
| 6 | 42 | 5.6 | 9793 | 3 | US-09-109-916-56 | Sequence 56, Appl |
| 7 | 42 | 5.6 | 9793 | 4 | US-09-886-156-56 | Sequence 56, Appl |
| 8 | 42 | 5.6 | 9793 | 4 | US-09-886-149-56 | Sequence 56, Appl |
| 9 | 42 | 5.6 | 9793 | 4 | US-09-886-150-56 | Sequence 56, Appl |
| 10 | 42 | 5.6 | 9793 | 4 | US-09-886-159-56 | Sequence 56, Appl |
| 11 | 41.8 | 5.6 | 1188 | 4 | US-09-328-352-2534 | Sequence 2534, Ap |
| 12 | 41.6 | 5.6 | 580073 | 4 | US-08-545-528D-1 | Sequence 1, Appl |
| 13 | 39.8 | 5.3 | 774 | 4 | US-09-134-001C-2190 | Sequence 2190, Ap |
| 14 | 39.8 | 5.3 | 6558 | 4 | US-09-491-356C-7 | Sequence 7, Appli |
| 15 | 38.8 | 5.2 | 1395 | 1 | US-07-991-867B-25 | Sequence 25, Appl |
| 16 | 38.8 | 5.2 | 1395 | 1 | US-08-107-755A-25 | Sequence 25, Appl |
| 17 | 38.8 | 5.2 | 1395 | 2 | US-08-544-332-25 | Sequence 25, Appl |
| 18 | 38.8 | 5.2 | 1395 | 4 | US-09-370-861A-25 | Sequence 25, Appl |
| 19 | 38.8 | 5.2 | 6768 | 1 | US-08-107-755A-1 | Sequence 1, Appl |
| 20 | 38.8 | 5.2 | 8457 | 1 | US-07-991-867B-1 | Sequence 1, Appl |
| 21 | 38.8 | 5.2 | 8457 | 1 | US-08-107-755A-1 | Sequence 1, Appl |
| 22 | 38.8 | 5.2 | 8457 | 2 | US-08-544-332-1 | Sequence 1, Appl |
| 23 | 38.6 | 5.2 | 1159 | 4 | US-09-181-585-1 | Sequence 1, Appli |
| 24 | 38.4 | 5.1 | 1159 | 3 | US-08-928-361B-2 | Sequence 2, Appli |
| 25 | 38.4 | 5.1 | 5511 | 3 | US-09-588-995A-2 | Sequence 2, Appli |
| 26 | 38.4 | 5.1 | 7334 | 3 | US-09-928-361B-1 | Sequence 1, Appli |
| 27 | 38.4 | 5.1 | 7334 | 4 | US-09-588-995A-1 | Sequence 1, Appli |
| 28 | 38.2 | 5.1 | 319608 | 4 | US-09-539-333D-1 | Sequence 1, Appli |
| 29 | 38.2 | 5.1 | 319608 | 4 | US-09-679-409-1. | Sequence 1, Appli |
| 30 | 38 | 5.1 | 1284 | 4 | US-09-134-001C-2306 | Sequence 2306, Ap |
| 31 | 37.8 | 5.1 | 580073 | 4 | US-08-545-528D-1 | Sequence 1, Appli |
| 32 | 37.6 | 5.0 | 9873 | 4 | US-09-328-352-1360 | Sequence 1360, Ap |
| 33 | 37.2 | 5.0 | 397 | 4 | US-09-253-691-3 | Sequence 3, Appli |
| 34 | 37 | 5.0 | 1037 | 4 | US-09-181-585-3 | Sequence 3, Appli |
| 35 | 37 | 5.0 | 1471 | 4 | US-09-181-585-2 | Sequence 2, Appl |
| 36 | 36.8 | 4.9 | 1086 | 1 | US-08-415-751-27 | Sequence 27, Appl |
| 37 | 36.8 | 4.9 | 1086 | 1 | US-08-700-651-1 | Sequence 1, Appl |
| 38 | 36.8 | 4.9 | 5163 | 3 | US-09-588-995A-4 | Sequence 4, Appli |
| 39 | 36.8 | 4.9 | 5163 | 4 | US-09-928-361B-4 | Sequence 4, Appli |
| 40 | 36.8 | 4.9 | 5163 | 4 | US-08-928-361B-4 | Sequence 4, Appli |
| 41 | 36.8 | 4.9 | 5318 | 3 | US-09-588-995A-3 | Sequence 3, Appli |
| 42 | 36.8 | 4.9 | 5318 | 4 | US-09-700-651-2 | Sequence 2, Appli |
| 43 | 36.8 | 4.9 | 5318 | 4 | US-08-928-361B-3 | Sequence 3, Appli |
| 44 | 36.8 | 4.9 | 1664976 | 4 | US-09-588-995A-3 | Sequence 3, Appli |
| 45 | 36.6 | 4.9 | 265 | 4 | US-09-491-356C-14 | Sequence 14, Appl |

ALIGNMENTS

```
RESULT 1
US-09-220-132-94
Sequence 94, Application US/09220132
Patent No. 6506607

GENERAL INFORMATION:
 APPLICANT:  Shyjan, Andrew W.
 TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
 TITLE OF INVENTION:  OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANC
 FILE REFERENCE:  07334-074001
 CURRENT APPLICATION NUMBER:  US/09/220,132
 CURRENT FILING DATE:  1998-12-23
 PRIOR APPLICATION NUMBER:  US 60/079,303
 PRIOR FILING DATE:  1998-03-25
 PRIOR APPLICATION NUMBER:  US 60/068,821
 PRIOR FILING DATE:  1997-12-24
 NUMBER OF SEQ ID NOS:  191
 SOFTWARE:  FastSEQ for Windows Version 4.0
SEQ ID NO 94
 LENGTH: 1497
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (1)..(1497)
  OTHER INFORMATION: n = A,T,C or G
US-09-220-132-94

Query Match             5.8%;  Score 43.2;  DB 4;  Length 1497;
Best Local Similarity  25.2%;  Pred. No. 0.037;
Matches  62;  Conservative  85;  Mismatches  98;  Indels  1;  Gaps  1;

Qy  112 GCTGCTTTATTAGCATCTACTGTATTTTCTTTGTTGAAAGAGATAGAGTTTCTGCAAAA 171
         ::!::  !:!  :  :!  ::! !::!!:!!:!:!!:!!::! !!:! ::::!:
Db  326 SYKRMWWKRWYWAAAWYCMWWKGKTTTTYMAAAWKGRAMYKGRAAARRGSMWTYCCRSMWMM 385

Qy  172 TGGAAAAACATCATTAACTGTAICTGGTCTTCGTTACTGGTATTGCTTTCTGGCATTACATG 231
         : !::!  !:! : ::  !:! :! : :!!:! !  : ! !::  !:! !::! !:::
Db  386 WTTWWWAMMMAWATTWAWTTTYMARKTYCMRGGSYTKRMMRGSCCYKGGGGSMWTYYWAR 445

Qy  232 TACATGAGAGGGGTATGGATTGAAAACTGGTGATTCGCCAACTGTATTTAGATACATTGAT 291
         :!: !:: !:: :!:! :!!!! :::::!! :!:   ! :!:: !!: ::!!
Db  446 SMARGRWTGSMWWAAWTKGGYMAWYYTGKKGRAARRKTKGGKWMYYTTKGYTYMAWTTTT 505

Qy  292 TGGTTACTAACAGTTCCTCTATTAATAATGTGAATTCTACTTAATTCTTGCTGCTGCAACT 351
         :!:! : :!! ! :!:: :!:! : !!:!  !:!!:!:!: :!!! :!:!
Db  506 WAWTWMCCMAWYYTYYYYKGATTGTAATT-ACCTCCAYTACTTYCATGATCCCCTACA 564

Qy  352 AATGTT 357
        ||||||
Db  565 AATATT 570
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: February 9, 2004, 16:16:59 ; Search time 341 Seconds
(without alignments)
8069.424 Million cell updates/sec

Title: US-09-847-513A-6
Perfect score: 747
Sequence: 1 atgaaattactcgatatt..........ttaaagaacttctaatgct 747
Scoring table: IDENTITY NUC
   Gapop 10.0 , Gapext 1.0
Searched: 2449703 seqs, 184181367 residues
   Listing first 45 summaries

Post-processing: Minimum Match 0%
   Maximum Match 100%

Total number of hits satisfying chosen parameters:    4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

## SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 747 | 100.0 | 747 | 11 | US-09-847-513A-6 | Sequence 6, Appli |
| c 2 | 747 | 100.0 | 105,184 | 11 | US-09-847-513A-36' | Sequence 1, Appli |
| 3 | 744 | 99.6 | 750 | 11 | US-09-847-513A-4 | Sequence 4, Appli |
| 4 | 739.2 | 99.0 | 750 | 11 | US-09-847-513A-32 | Sequence 32, Appl |
| 5 | 737.6 | 98.7 | 750 | 11 | US-09-847-513A-10 | Sequence 10, Appl |
| 6 | 728 | 97.5 | 750 | 11 | US-09-847-513A-38 | Sequence 38, Appl |
| 7 | 728 | 97.5 | 750 | 11 | US-09-847-513A-44 | Sequence 44, Appl |
| 8 | 713.6 | 95.5 | 750 | 11 | US-09-847-513A-34 | Sequence 34, Appl |
| 9 | 705.6 | 94.5 | 750 | 11 | US-09-847-513A-30 | Sequence 30, Appl |
| 10 | 705.6 | 94.5 | 750 | 11 | US-09-847-513A-42 | Sequence 42, Appl |
| 11 | 705 | 94.2 | 750 | 11 | US-09-847-513A-28 | Sequence 28, Appl |
| 12 | 704 | 94.2 | 750 | 11 | US-09-847-513A-40 | Sequence 40, Appl |
| 13 | 700.8 | 93.8 | 750 | 11 | US-09-847-513A-12 | Sequence 12, Appl |
| 14 | 697.6 | 93.4 | 750 | 11 | US-09-847-513A-46 | Sequence 46, Appl |
| 15 | 696 | 93.2 | 750 | 11 | US-09-847-513A-8 | Sequence 8, Appli |
| 16 | 694.4 | 93.0 | 750 | 11 | US-09-847-513A-6 | Sequence 14, Appl |
| 17 | 694.4 | 93.0 | 750 | 11 | US-09-847-513A-36 | Sequence 36, Appl |
| 18 | 689.6 | 92.3 | 750 | 11 | US-09-847-513A-24 | Sequence 24, Appl |
| 19 | 688 | 92.1 | 750 | 11 | US-09-847-513A-26 | Sequence 26, Appl |
| 20 | 459 | 61.4 | 753 | 11 | US-09-847-513A-18 | Sequence 18, Appl |
| 21 | 459 | 61.4 | 753 | 11 | US-09-847-513A-62 | Sequence 62, Appl |
| 22 | 457.4 | 61.2 | 753 | 11 | US-09-847-513A-48 | Sequence 48, Appl |
| 23 | 457.4 | 61.2 | 753 | 11 | US-09-847-513A-56 | Sequence 56, Appl |
| 24 | 455.8 | 61.0 | 753 | 11 | US-09-847-513A-22 | Sequence 22, Appl |
| 25 | 455.8 | 61.0 | 753 | 11 | US-09-847-513A-52 | Sequence 52, Appl |
| 26 | 454.2 | 60.8 | 753 | 11 | US-09-847-513A-54 | Sequence 54, Appl |
| 27 | 452.6 | 60.6 | 753 | 11 | US-09-847-513A-20 | Sequence 20, Appl |
| 28 | 452.6 | 60.6 | 753 | 11 | US-09-847-513A-50 | Sequence 50, Appl |
| 29 | 452.6 | 60.6 | 753 | 11 | US-09-847-513A-58 | Sequence 58, Appl |
| 30 | 452.6 | 60.6 | 753 | 11 | US-09-847-513A-64 | Sequence 64, Appl |
| 31 | 451 | 60.4 | 753 | 11 | US-09-847-513A-60 | Sequence 60, Appl |
| 32 | 449.4 | 60.2 | 753 | 11 | US-09-847-513A-16 | Sequence 16, Appl |
| 33 | 48.2 | 6.5 | 3673778 | 13 | US-10-312-841-2 | Sequence 2, Appli |
| 34 | 47.4 | 6.3 | 7758 | 13 | US-10-311-455-1076 | Sequence 1076, Ap |
| 35 | 45.4 | 6.1 | 99116 | 13 | US-10-298-122-1 | Sequence 1, Appli |
| c 36 | 45.2 | 6.1 | 3673778 | 13 | US-10-312-841-1 | Sequence 1, Appli |
| 37 | 45 | 6.0 | 11155 | 13 | US-10-311-455-1916 | Sequence 1916, Ap |
| 38 | 45 | 6.0 | 12138 | 13 | US-10-240-453-210 | Sequence 210, App |
| 39 | 45 | 6.0 | 12138 | 13 | US-10-311-455-578 | Sequence 578, App |
| 40 | 44.8 | 6.0 | 8297 | 13 | US-10-311-455-2047 | Sequence 2047, Ap |
| 41 | 44.6 | 6.0 | 6522 | 13 | US-10-311-455-1023 | Sequence 1023, Ap |
| 42 | 44.6 | 6.0 | 9652 | 13 | US-10-311-455-882 | Sequence 882, App |
| 43 | 44.4 | 5.9 | 14708 | 13 | US-10-311-455-2217 | Sequence 2217, Ap |
| 44 | 44.4 | 5.9 | 14708 | 13 | US-10-311-455-3323 | Sequence 3323, Ap |
| 45 | 44.4 | 5.9 | 14708 | 15 | US-10-229-676-221 | Sequence 221, App |

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## ALIGNMENTS

RESULT 1
US-09-847-513A-6
; Sequence 6, Application US/09847513A
; Publication No. US20030104375A1
GENERAL INFORMATION:
; APPLICANT: MBARI
; APPLICANT: DeLong, Edward
; APPLICANT: Beja, Oded
; TITLE OF INVENTION: Light-driven energy generation using proteorhodopsin
; FILE REFERENCE: MBA-101
; CURRENT APPLICATION NUMBER: US/09/847,513A
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: 60/201,602
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 6
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Naturally occurring gamma proteobacterium
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(747)
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Beja,O., Aravind,L., Koonin,E.V., Suzuki,M.T., Hadd,A., Nguyen,L.P.,
; AUTHORS: Jovanovich,S.B., Gates,C.M., Feldman,R.A., Spudich,J.L., Spudich,E.N. and
; TITLE: Bacterial rhodopsin: evidence for a new type of phototrophy in the sea
; JOURNAL: Science
; VOLUME: 289
; ISSUE: 5486
; PAGES: 1902-1906
; DATE: 2000-09-15
; DATABASE ACCESSION NUMBER: AAG10475
; DATABASE ENTRY DATE: 2000-06-15
; RELEVANT RESIDUES: (1)..(747)
US-09-847-513A-6

GenCore version 5.1.6
Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        February  9, 2004, 15:20:39 ; Search time 2081 Seconds
                                              (without alignments)
                                              8724.384 Million cell updates/sec

Title:         US-09-847-513A-6
Perfect score: 747
Sequence:      1 atgaaattattactgatatt..........ttaagaatcttctaatgct 747

Scoring table: IDENTITY NUC
               Gapop 10.0 , Gapext 1.0

Searched:      22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters:      45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     EST:*
               1: em_estba:*
               2: em_esthum:*
               3: em_estin:*
               4: em_estmu:*
               5: em_estov:*
               6: em_estpl:*
               7: em_estro:*
               8: em_htc:*
               9: em_esti:*
              10: gb_est2:*
              11: gb_htc:*
              12: gb_est3:*
              13: gb_est4:*
              14: gb_est5:*
              15: em_estfun:*
              16: em_estom:*
              17: em_gss_hum:*
              18: em_gss_inv:*
              19: em_gss_pln:*
              20: em_gss_vrt:*
              21: em_gss_fun:*
              22: em_gss_mam:*
              23: em_gss_mus:*
              24: em_gss_pro:*
              25: em_gss_rod:*
              26: em_gss_phg:*
              27: em_gss_vrl:*
              28: gb_gss1:*
              29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 63.2 | 8.5 | 1101 | 29 | CNS00LT2 | AL078714 Drosophil |
| 2 | 58.2 | 7.7 | 1101 | 29 | CNS002J2 | AL063072 Drosophil |
| 3 | 56.4 | 7.6 | 712 | 13 | BX416727 | BX416727 Drosophil |
| c 4 | 52.6 | 7.0 | 1101 | 29 | CNS00L00 | AL068607 Drosophil |

ALIGNMENTS

RESULT 1
CNS00LT2
LOCUS        CNS00LT2                 1101 bp    DNA     linear   GSS 14-JUN-1999
DEFINITION   Drosophila melanogaster genome survey sequence TET3 end of BAC:
             BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit
             fly), genomic survey sequence.
ACCESSION    AL078714
VERSION      AL078714.1  GI:5102004
KEYWORDS     GSS.
SOURCE       Drosophila melanogaster (fruit fly)
  ORGANISM   Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
             Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
             Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1  (bases 1 to 1101)
  AUTHORS    Genoscope.
  TITLE      Direct Submission
  JOURNAL    Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
             collaboration with the Berkeley Drosophila Genome Project (BDGP).
             The BDGP is constructing a physical map of the Drosophila
             melanogaster genome using these BACs. For further information
             please see http://www.fruitfly.org The BDGP Drosophila
             melanogaster BAC library was prepared by Kazutoyo Oseegawa and
             Aaron Mammoser in Pieter de Jong's laboratory in the Department of

| c | 5 | 51.8 | 6.9 | 610 | 12 | CNS07BBR | BJ334069 BJ334069 |
| | 6 | 51.8 | 6.9 | 1180 | 13 | CNS0010OX | BX436369 BX436369 |
| | 7 | 51.4 | 6.9 | 885 | 13 | CNS0010lX | BX425603 BX425603 |
| c | 8 | 51.2 | 6.9 | 1101 | 29 | CNS0039G | AL063921 Drosophil |
| | 9 | 51 | 6.8 | 1039 | 13 | CNS060OI | BX378025 BX378025 |
| | 10 | 50.8 | 6.8 | 990 | 29 | CNS0100X | AL065624 Drosophil |
| | 11 | 50.4 | 6.7 | 1101 | 29 | CNS00609 | AL098379 Drosophil |
| c | 12 | 50.2 | 6.7 | 964 | 29 | CNS07E8R | AL441457 T7 end of |
| | 13 | 49.4 | 6.6 | 354 | 28 | AQ903139 | AQ903139 GSSTC0729 |
| | 14 | 49 | 6.6 | 818 | 29 | B2229437 | B2229437 CH230-401 |
| c | 15 | 49 | 6.6 | 909 | 13 | CNS012CM | BX391298 BX391298 |
| | 16 | 49 | 6.5 | 1124 | 13 | BX436282 | BX436282 AGENCOURT |
| | 17 | 48.8 | 6.5 | 1097 | 14 | CD048504 | CD048504 AGENCOURT |
| c | 18 | 48.8 | 6.5 | 1102 | 9 | AL572008 | AL572008 AL572008 |
| | 19 | 48.8 | 6.5 | 1201 | 13 | BX403984 | BX403984 BX403984 |
| c | 20 | 48.8 | 6.5 | 1084 | 29 | CC220905 | CC220905 CH261-71A |
| | 21 | 48.2 | 6.5 | 699 | 12 | BM630395 | BM630395 170006875 |
| c | 22 | 48 | 6.4 | 423 | 9 | AU034896 | AU034896 AU034896 |
| | 23 | 48 | 6.4 | 854 | 29 | CNS012CM | AL101392 Drosophil |
| c | 24 | 47.6 | 6.4 | 966 | 29 | B2695402 | SP  Ba006 |
| | 25 | 47.6 | 6.4 | 1198 | 12 | BM452445 | BM452445 AGENCOURT |
| c | 26 | 47.6 | 6.3 | 690 | 29 | CNS008BY | AL051480 Drosophil |
| | 27 | 47.4 | 6.3 | 767 | 29 | CNS00AQX | AL055924 Drosophil |
| c | 28 | 47.2 | 6.3 | 956 | 13 | CNS00LCP | AL067950 Drosophil |
| | 29 | 47.2 | 6.3 | 1098 | 29 | BX377526 | BX377526 BX377526 |
| c | 30 | 47.2 | 6.3 | 1101 | 29 | CNS017QV | AL108385 Drosophil |
| | 31 | 47 | 6.3 | 797 | 12 | BJ324597 | BJ324597 BJ324597 |
| c | 32 | 47 | 6.3 | 1007 | 29 | CNS06X9S | AL419462 T3 end of |
| | 33 | 46.8 | 6.3 | 527 | 9 | AU271052 | AU271052 AU271052 |
| c | 34 | 46.8 | 6.3 | 548 | 9 | AU271053 | AU271053 AU271053 |
| | 35 | 46.8 | 6.3 | 992 | 29 | CNS002SH | AL063310 Drosophil |
| c | 36 | 46.8 | 6.2 | 1001 | 29 | CNS0155H | AL105023 Drosophil |
| | 37 | 46.6 | 6.2 | 1200 | 9 | AL548181 | AL548181 AL548181 |
| c | 38 | 46.2 | 6.2 | 711 | 28 | B2098560 | B2098560 CH230-235 |
| | 39 | 46.2 | 6.2 | 1163 | 13 | BX415221 | BX415221 BX415221 |
| c | 40 | 46.2 | 6.2 | 1167 | 29 | CNS07360 | AL427102 clone BA0 |
| | 41 | 46.2 | 6.2 | 1175 | 12 | BI872945 | BI872945 603398115 |
| c | 42 | 46.2 | 6.2 | 1201 | 13 | BX336382 | BX336382 BX336382 |
| c | 43 | 46 | 6.2 | 1201 | 13 | BX417935 | BX417935 BX417935 |
| c | 44 | 45.8 | 6.1 | 884 | 29 | CNS0129A | AL101272 Drosophil |
| | 45 | 45.6 | 6.1 | 423 | 28 | A2392278 | A2392278 1M0154M12 |

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:          May 27, 2003, 04:11:46 ; Search time 382.939 Seconds
                 (without alignments)
                 4410.621 Million cell updates/sec

Title:
Perfect score:   750
Sequence:        1 atgggtaaattattactgat........ttaaagaatcttctaatgct 750
                 US-09-847-513A-4

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 100 summaries

Database :       N_Geneseq_101002:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 750 | 100.0 | 750 | 24 | ABK24125 | DNA encoding Prote |
| 2 | 745.2 | 99.4 | 750 | 24 | ABK24139 | DNA encoding Prote |
| 3 | 744 | 99.2 | 747 | 24 | ABK24126 | DNA encoding Prote |
| 4 | 744 | 99.2 | 105184 | 24 | ABK24122 | Bacterial artifici |
| 5 | 743.6 | 99.1 | 750 | 24 | ABK24128 | DNA encoding Prote |
| 6 | 734 | 97.9 | 750 | 24 | ABK24142 | DNA encoding Prote |
| 7 | 734 | 97.9 | 750 | 24 | ABK24145 | DNA encoding Prote |
| 8 | 719.6 | 95.9 | 750 | 24 | ABK24140 | DNA encoding Prote |
| 9 | 711.6 | 94.9 | 750 | 24 | ABK24138 | DNA encoding Prote |
| 10 | 711.6 | 94.9 | 750 | 24 | ABK24144 | DNA encoding Prote |
| 11 | 710 | 94.7 | 750 | 24 | ABK24137 | DNA encoding Prote |
| 12 | 706.8 | 94.2 | 750 | 24 | ABK24143 | DNA encoding Prote |
| 13 | 703.6 | 93.8 | 750 | 24 | ABK24129 | DNA encoding Prote |
| 14 | 702 | 93.6 | 750 | 24 | ABK24146 | DNA encoding Prote |
| 15 | 700.4 | 93.4 | 750 | 24 | ABK24127 | DNA encoding Prote |
| 16 | 700.4 | 93.4 | 750 | 24 | ABK24130 | DNA encoding Prote |
| 17 | 700.4 | 93.4 | 750 | 24 | ABK24141 | DNA encoding Prote |
| 18 | 695.6 | 92.7 | 750 | 24 | ABK24136 | DNA encoding Prote |
| 19 | 694 | 92.5 | 750 | 24 | ABK24135 | DNA encoding Prote |
| 20 | 465 | 62.0 | 753 | 24 | ABK24132 | DNA encoding Prote |
| 21 | 465 | 62.0 | 753 | 24 | ABK24154 | DNA encoding Prote |
| 22 | 463.4 | 61.8 | 753 | 24 | ABK24147 | DNA encoding Prote |
| 23 | 461.8 | 61.6 | 753 | 24 | ABK24151 | DNA encoding Prote |
| 24 | 460.2 | 61.4 | 753 | 24 | ABK24134 | DNA encoding Prote |
| 25 | 458.6 | 61.1 | 753 | 24 | ABK24133 | DNA encoding Prote |
| 26 | 458.6 | 61.1 | 753 | 24 | ABK24150 | DNA encoding Prote |
| 27 | 458.6 | 61.1 | 753 | 24 | ABK24149 | DNA encoding Prote |
| 28 | 458.6 | 61.1 | 753 | 24 | ABK24148 | DNA encoding Prote |
| 29 | 457 | 60.9 | 753 | 24 | ABK24152 | DNA encoding Prote |
| 30 | 455.4 | 60.7 | 753 | 24 | ABK24155 | DNA encoding Prote |
| 31 | 457 | 60.9 | 753 | 24 | ABK24153 | DNA encoding Prote |
| 32 | 455.4 | 60.7 | 753 | 24 | ABK24131 | DNA encoding Prote |
| 33 | 48.6 | 6.5 | 6767 | 22 | AAS46608 | Tumour suppressor |
| 34 | 48.6 | 6.5 | 6317 | 22 | ABL70427 | Chemically treated |
| 35 | 48.6 | 6.5 | 7131 | 24 | AAS61360 | Human gene regulat |
| 36 | 47.4 | 6.3 | 7131 | 24 | ABK31450 | Signal transductio |
| 37 | 45.6 | 6.1 | 7758 | 24 | ABK33103 | Human immune syste |
| 38 | 45.6 | 6.1 | 6317 | 24 | ABL49312 | Human polynucleoti |
| 39 | 45.6 | 6.1 | 6317 | 24 | ABL32409 | Human immune syste |
| 40 | 45 | 6.0 | 11155 | 24 | ABL32605 | Human immune syste |
| 41 | 45 | 6.0 | 12138 | 24 | ABL33943 | Human immune syste |
| 42 | 44.8 | 6.0 | 12138 | 24 | ABK28336 | DNA transcription |
| 43 | 44.8 | 6.0 | 6419 | 24 | ABL32266 | Human immune syste |
| 44 | 44.8 | 6.0 | 8297 | 24 | ABL34074 | Human immune syste |
| 45 | 45.6 | 6.1 | 56153 | 24 | AAP47794 | Tumour suppressor |
| 46 | 44.6 | 5.9 | 390 | 24 | ABA97217 | Ribosomal protein |
| 47 | 44.6 | 5.9 | 6522 | 24 | ABL33050 | Tumour suppressor |
| 48 | 44.6 | 5.9 | 9652 | 24 | ABL32909 | Human immune syste |
| 49 | 44.6 | 5.9 | 12639 | 24 | ABN80107 | Human chemically m |
| 50 | 44.6 | 5.9 | 7131 | 24 | ABL07253 | Drosophila melanog |
| 51 | 44.4 | 5.9 | 6030 | 23 | AAS45456 | Human DNA for stag |
| 52 | 44.4 | 5.9 | 6030 | 22 | AAS61348 | Drosophila melanog |
| 53 | 44.4 | 5.9 | 6030 | 24 | ABK28312 | Human angiogenesis |
| 54 | 44.4 | 5.9 | 8776 | 24 | ABK40068 | Chemically pretrea |
| 55 | 44.4 | 5.9 | 13531 | 23 | ABL07252 | Human gene regulat |
| 56 | 44.4 | 5.9 | 12000 | 24 | ABL32000 | DNA transcription |
| 57 | 44.2 | 5.9 | 14362 | 24 | ABK34014 | Human chemically p |
| 58 | 44.2 | 5.9 | 14708 | 22 | ABK67047 | Drosophila melanog |
| 59 | 44.2 | 5.9 | 14708 | 22 | AAS45513 | Human chemically p |
| 60 | 44.2 | 5.9 | 14708 | 22 | AAS46773 | Human DNA for stag |
| 61 | 44.2 | 5.9 | 14708 | 22 | AAS62324 | Human immune syste |
| 62 | 44.2 | 5.9 | 14708 | 24 | ABL34244 | Chemically pretrea |
| 63 | 43.8 | 5.8 | 14708 | 24 | ABL32603 | Chemically pretrea |
| 64 | 43.8 | 5.8 | 6292 | 22 | ABL28449 | Tumour suppressor |
| 65 | 43.2 | 5.8 | 7001 | 24 | AAS46735 | Human immune syste |
| 66 | 43 | 5.7 | 4122 | 23 | ABK33919 | DNA transcription |
| 67 | 43 | 5.7 | 5273 | 24 | ABL28439 | Human immune syste |
| 68 | 43 | 5.7 | 5535 | 24 | ABL32875 | Tumour suppressor |
| 69 | 43 | 5.7 | 6708 | 23 | ABL32936 | Human immune syste |
| 70 | 43 | 5.7 | 10266 | 17 | ABL28438 | Drosophila melanog |
| 71 | 43 | 5.7 | 12000 | 24 | AAT33007 | Human chemically p |
| 72 | 43 | 5.7 | 11178 | 24 | ABL70507 | Mouse SRY-related |
| 73 | 43 | 5.7 | 11178 | 24 | AAS61059 | Chemically treated |
| 74 | 43 | 5.7 | 11178 | 24 | ABK31172 | Human gene regulat |
| 75 | 42.8 | 5.7 | 14704 | 13 | AAQ20685 | Signal transductio |
| 76 | 42.8 | 5.7 | 875 | 22 | AAI95044 | PRS 741 insert con |
| 77 | 42.6 | 5.7 | 5649 | 24 | AAS46384 | Human neuroblastom |
| 78 | 42.6 | 5.7 | 5649 | 24 | ABK40008 | Tumour suppressor |
| 79 | 42.6 | 5.7 | 5649 | 24 | ABL32849 | Human chemically p |
| 80 | 42.4 | 5.7 | 6283 | 24 | ABL32088 | Human immune syste |
| 81 | 42.2 | 5.6 | 16688 | 22 | AAS46556 | Human immune syste |
| 82 | 42.2 | 5.6 | 1448 | 24 | ABQ14794 | Tumour suppressor |
|  | 42.2 | 5.6 | 1448 | 24 | ABQ14795 | Oligonucleotide fo |

Listing first 100 summaries

1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.6

Run on: May 27, 2003, 05:27:32 ; Search time 2805.73 Seconds
(without alignments)
7779.478 Million cell updates/sec

Title: US-09-847-513A-4

Perfect score: 750

Sequence: 1 atgggtaaattattactgat.........ttaaagaatcttctaatgct 750

Scoring table: IDENTITY_NUC

Gapop 10.0 / Gapext 1.0

Searched: 2054640 seqs, 1455102878 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : GenEmbl:*

```
 1: gb_ba:*              21: em_or:*
 2: gb_htg:*             22: em_ov:*
 3: gb_in:*              23: em_pat:*
 4: gb_om:*              24: em_ph:*
 5: gb_ov:*              25: em_pl:*
 6: gb_pat:*             26: em_ro:*
 7: gb_ph:*              27: em_sts:*
 8: gb_pl:*              28: em_un:*
 9: gb_pr:*              29: em_vi:*
10: gb_ro:*              30: em_hum:*
11: gb_sts:*             31: em_htg_inv:*
12: gb_sy:*              32: em_htg_other:*
13: gb_un:*              33: em_htg_mus:*
14: gb_vi:*              34: em_htg_pln:*
15: em_ba:*              35: em_htg_rod:*
16: em_fun:*             36: em_htg_mam:*
17: em_hum:*             37: em_htg_vrt:*
18: em_in:*              38: em_sy:*
19: em_mu:*              39: em_htgo_hum:*
20: em_om:*              40: em_htgo_mus:*
                        41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| 1 | c | 744 | 99.2 | 105184 | 1 | AF279106 | AF279106 Uncultured |
| 2 | | 740.2 | 98.7 | 748 | 1 | AF349987 | AF349987 Uncultured |
| 3 | | 734 | 97.9 | 753 | 1 | AF349990 | AF349990 Uncultured |
| 4 | | 734 | 97.9 | 753 | 1 | AF349993 | AF349993 Uncultured |
| 5 | | 719.6 | 95.9 | 753 | 1 | AF349988 | AF349988 Uncultured |
| 6 | | 711.6 | 94.9 | 753 | 1 | AF349986 | AF349986 Uncultured |
| 7 | | 711.6 | 94.9 | 753 | 1 | AF349992 | AF349992 Uncultured |
| 8 | | 710 | 94.7 | 753 | 1 | AF349985 | AF349985 Uncultured |
| 9 | | 706.8 | 94.2 | 753 | 1 | AF349991 | AF349991 Uncultured |
| 10 | | 703.6 | 93.8 | 757 | 1 | AF349977 | AF349977 Uncultured |
| 11 | | 702 | 93.6 | 753 | 1 | AF349994 | AF349994 Uncultured |
| 12 | | 700.4 | 93.4 | 753 | 1 | AF349976 | AF349976 Uncultured |
| 13 | | 700.4 | 93.4 | 753 | 1 | AF349978 | AF349978 Uncultured |
| 14 | | 695.6 | 92.7 | 753 | 1 | AF349989 | AF349989 Uncultured |
| 15 | | 695.4 | 92.7 | 748 | 1 | AF349984 | AF349984 Uncultured |
| 16 | | 694 | 92.5 | 753 | 1 | AF349983 | AF349983 Uncultured |
| 17 | | 465 | 62.0 | 756 | 1 | AF350002 | AF350002 Uncultured |
| 18 | | 465 | 62.0 | 757 | 1 | AF349980 | AF349980 Uncultured |
| 19 | | 463.4 | 61.8 | 756 | 1 | AF349995 | AF349995 Uncultured |
| 20 | | 463.4 | 61.8 | 756 | 1 | AF349998 | AF349998 Uncultured |
| 21 | | 461.8 | 61.6 | 756 | 1 | AF349982 | AF349982 Uncultured |
| 22 | | 461.8 | 61.6 | 756 | 1 | AF349997 | AF349997 Uncultured |
| 23 | | 460.2 | 61.4 | 756 | 1 | AF349999 | AF349999 Uncultured |
| 24 | | 458.6 | 61.1 | 756 | 1 | AF349996 | AF349996 Uncultured |
| 25 | | 458.6 | 61.1 | 756 | 1 | AF349981 | AF349981 Uncultured |
| 26 | | 458.6 | 61.1 | 756 | 1 | AF350000 | AF350000 Uncultured |
| 27 | | 457 | 60.9 | 756 | 1 | AF350003 | AF350003 Uncultured |
| 28 | | 457 | 60.9 | 756 | 1 | AF350001 | AF350001 Uncultured |
| 29 | | 455.4 | 60.7 | 756 | 1 | AF349979 | AF349979 Uncultured |
| 30 | | 455.4 | 60.7 | 756 | 1 | AF096320 | AF096320 Uncultured |
| 31 | c | 54.4 | 7.6 | 235542 | 2 | AC115666 | AC115666 Rattus no |
| 32 | | 57.2 | 7.3 | 169163 | 2 | AC112091 | AC112091 Rattus no |
| 33 | | 54 | 7.2 | 84472 | 2 | AC096684 | AC096684 Takifugu |
| 34 | | 54 | 7.2 | 169163 | 2 | AC102515 | AC102515 Rattus no |
| 35 | | 53.6 | 7.1 | 163034 | 2 | AC099432 | AC099432 Rattus no |
| 36 | | 53.4 | 7.1 | 108476 | 2 | AC094507 | AC094507 Mus muscu |
| 37 | | 52.8 | 7.0 | 138156 | 2 | AC111482 | AC111482 Rattus no |
| 38 | c | 52.6 | 7.0 | 131346 | 2 | AC119558 | AC119558 Rattus no |
| 39 | | 52.6 | 7.0 | 167439 | 2 | AC102515 | AC102515 Mus muscu |
| 40 | c | 52.6 | 7.0 | 175699 | 2 | AC129136 | AC129136 Mus muscu |
| 41 | | 52.6 | 7.0 | 201097 | 2 | AC112091 | AC112091 Rattus no |
| 42 | c | 52.6 | 7.0 | 207364 | 2 | AL671857 | AL671857 Mus muscu |
| 43 | | 52 | 6.9 | 125026 | 10 | AC091616 | AC091616 Dictyoste |
| 44 | c | 52 | 6.9 | 165337 | 2 | AC096212 | AC096212 Rattus no |
| 45 | | 51.6 | 6.9 | 141008 | 2 | AC126897 | AC126897 Rattus no |
| 46 | c | 51.6 | 6.9 | 264522 | 2 | AC090437 | AC090437 Mus muscu |
| 47 | | 51.4 | 6.9 | 7921 | 3 | AF153362 | AF153362 Dictyoste |
| 48 | | 51.2 | 6.8 | 138825 | 2 | AC125755 | AC125755 Rattus no |
| 49 | | 51 | 6.8 | 97683 | 2 | AC096844 | AC096844 Rattus no |
| 50 | c | 50.4 | 6.7 | 144225 | 2 | AC116548 | AC116548 Dictyoste |
| 51 | | 50.4 | 6.7 | 151705 | 2 | AC118320 | AC118320 Rattus no |
| 52 | c | 50.2 | 6.7 | 151705 | 2 | AC117900 | AC117900 Rattus no |
| 53 | | 50.2 | 6.7 | 173540 | 2 | AC119697 | AC119697 Rattus no |
| 54 | | 49.8 | 6.6 | 196650 | 2 | AC111217 | AC111217 Rattus no |
| 55 | | 49.6 | 6.6 | 174599 | 2 | AC106523 | AC106523 Takifugu |
| 56 | c | 49.4 | 6.6 | 71553 | 2 | AC096997 | AC096997 Dictyoste |
| 57 | | 49 | 6.5 | 156533 | 2 | AC117070 | AC117070 Rattus no |
| 58 | | 49 | 6.5 | 97059 | 2 | AC129683 | AC129683 Rattus no |
| 59 | | 49 | 6.5 | 101739 | 2 | AC114251 | AC114251 Rattus no |
| 60 | c | 49 | 6.5 | 161799 | 2 | AC089891 | AC089891 Rattus no |
| 61 | | 49 | 6.5 | 197627 | 2 | AC112735 | AC112735 Rattus no |
| 62 | c | 48.8 | 6.5 | 22526 | 2 | AC130161 | AC130161 Rattus no |
| 63 | | 48.8 | 6.5 | 79981 | 2 | AC129738 | AC129738 Rattus no |
| 64 | | 48.8 | 6.5 | 127573 | 2 | AC097158 | AC097158 Rattus no |
| 65 | c | 48.8 | 6.5 | 165318 | 2 | AC126747 | AC126747 Rattus no |

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:            May 27, 2003, 06:04:58 ; Search time 79.3898 Seconds
                   (without alignments)
                   2897.193 Million cell updates/sec

Title:
Perfect score:    750
Sequence:          1 atgggtaaattattaactgat.........ttaaagaatcttctaatgct 750
                   US-09-847-513A-4

Scoring table:     IDENTITY_NUC
                   Gapop 10.0, Gapext 1.0

Searched:          444362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:    882724

Post-processing:   Minimum Match 0%
                   Maximum Match 100%
                   Listing first 100 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :  1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
               Issued_Patents_NA:*
            2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
            3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
            4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
            5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result NO. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| c 1 | 42 | 5.6 | 9793 | 1 | US-08-470-202-56 | Sequence 56, Appl |
| c 2 | 42 | 5.6 | 9793 | 1 | US-08-471-770-56 | Sequence 56, Appl |
| c 3 | 42 | 5.6 | 9793 | 2 | US-08-468-059-56 | Sequence 56, Appl |
| c 4 | 42 | 5.6 | 9793 | 1 | US-09-109-916-56 | Sequence 56, Appl |
| c 5 | 39.8 | 5.3 | 774 | 4 | US-09-134-001C-2190 | Sequence 2190, Ap |
| c 6 | 38.8 | 5.2 | 1395 | 1 | US-07-991-867B-25 | Sequence 25, Appl |
| c 7 | 38.8 | 5.2 | 1395 | 1 | US-08-107-755A-25 | Sequence 25, Appl |
| c 8 | 38.8 | 5.2 | 1395 | 2 | US-08-544-332-25 | Sequence 25, Appl |
| c 9 | 38.8 | 5.2 | 1395 | 4 | US-09-370-861A-25 | Sequence 25, Appl |
| 10 | 38.8 | 5.2 | 6768 | 1 | US-08-107-755A-1 | Sequence 1, Appl |
| 11 | 38.8 | 5.2 | 8457 | 1 | US-07-991-867B-1 | Sequence 1, Appl |
| 12 | 38.8 | 5.2 | 8457 | 1 | US-08-107-755A-1 | Sequence 1, Appl |
| 13 | 38.8 | 5.2 | 8457 | 2 | US-08-544-332-1 | Sequence 1, Appl |
| 14 | 38.4 | 5.1 | 8457 | 4 | US-09-370-861A-1 | Sequence 1, Appl |
| 15 | 38.4 | 5.1 | 7334 | 3 | US-08-928-361B-1 | Sequence 1, Appl |
| 16 | 38 | 5.1 | 1284 | 4 | US-09-134-001C-2306 | Sequence 2306, Ap |
| c 17 | 37.2 | 5.0 | 397 | 3 | US-09-253-691-3 | Sequence 3, Appl |
| c 18 | 36.8 | 4.9 | 1086 | 1 | US-08-415-751-27 | Sequence 27, Appl |
| c 19 | 36.8 | 4.9 | 1086 | 1 | US-08-415-751-28 | Sequence 28, Appl |
| c 20 | 36.8 | 4.9 | 5163 | 3 | US-08-700-651-1 | Sequence 1, Appl |
| c 21 | 36.8 | 4.9 | 5163 | 3 | US-08-700-651-2 | Sequence 2, Appl |
| c 22 | 36.8 | 4.9 | 5318 | 3 | US-08-928-361B-4 | Sequence 4, Appl |
| c 23 | 36.8 | 4.9 | 1447 | 3 | US-09-227-357-17 | Sequence 17, Appl |
| c 24 | 36.6 | 4.9 | 7218 | 2 | US-08-232-463-14 | Sequence 14, Appl |
| c 25 | 36.4 | 4.9 | 5318 | 3 | US-08-928-361B-3 | Sequence 3, Appl |
| c 26 | 36 | 4.8 | 2190 | 4 | US-09-625-188-19 | Sequence 19, Appl |
| c 27 | 35.6 | 4.7 | 663 | 4 | US-08-998-416-187 | Sequence 187, App |
| c 28 | 35.6 | 4.7 | 719 | 4 | US-08-998-416-1138 | Sequence 1138, Ap |
| c 29 | 35.6 | 4.7 | 856 | 4 | US-08-998-416-289 | Sequence 289, App |
| c 30 | 35.4 | 4.7 | 11091 | 5 | US-09-134-001C-2243 | Sequence 2243, Ap |
| c 31 | 35.4 | 4.7 | 2824 | 4 | US-07-757-022B-13 | Sequence 13, Appl |
| c 32 | 35.2 | 4.7 | 3065 | 4 | US-07-757-022B-83 | Sequence 83, Appl |
| c 33 | 35.2 | 4.7 | 3117 | 4 | US-07-757-022B-73 | Sequence 73, Appl |
| c 34 | 35.2 | 4.7 | 3148 | 4 | US-07-814-964-12 | Sequence 12, Appl |
| c 35 | 35.2 | 4.7 | 3292 | 1 | US-08-258-442-12 | Sequence 12, Appl |
| c 36 | 35.2 | 4.7 | 3292 | 5 | US-07-757-022B-57 | Sequence 57, Appl |
| c 37 | 35.2 | 4.7 | 3292 | 1 | PCT-US92-11107-12 | Sequence 12, Appl |
| c 38 | 35.2 | 4.7 | 3376 | 3 | US-08-320-559-29 | Sequence 29, Appl |
| c 39 | 35.2 | 4.7 | 3376 | 4 | US-08-545-860D-29 | Sequence 29, Appl |
| c 40 | 35.2 | 4.7 | 3376 | 5 | PCT-US94-04496-29 | Sequence 29, Appl |
| c 41 | 35.2 | 4.7 | 3420 | 5 | PCT-US95-10668-29 | Sequence 29, Appl |
| c 42 | 35.2 | 4.7 | 3420 | 5 | US-07-757-022B-103 | Sequence 103, App |
| c 43 | 35.2 | 4.7 | 3831 | 4 | US-07-757-022B-43 | Sequence 43, Appl |
| c 44 | 35.2 | 4.7 | 3831 | 4 | US-08-961-527-291 | Sequence 291, App |
| c 45 | 35.2 | 4.7 | 3936 | 4 | US-07-757-022B-41 | Sequence 41, Appl |
| c 46 | 35.2 | 4.7 | 3942 | 4 | US-07-757-022B-141 | Sequence 141, App |
| c 47 | 35.2 | 4.7 | 3945 | 4 | US-07-757-022B-49 | Sequence 49, Appl |
| c 48 | 35.2 | 4.7 | 3963 | 4 | US-07-757-022B-45 | Sequence 45, Appl |
| c 49 | 35.2 | 4.7 | 3963 | 4 | US-07-757-022B-59 | Sequence 59, Appl |
| c 50 | 35.2 | 4.7 | 4065 | 4 | US-07-757-022B-47 | Sequence 47, Appl |
| c 51 | 35.2 | 4.7 | 4092 | 4 | US-07-757-022B-39 | Sequence 39, Appl |
| c 52 | 35.2 | 4.7 | 4215 | 4 | US-07-757-022B-51 | Sequence 51, Appl |
| c 53 | 35.2 | 4.7 | 4086 | 4 | US-07-757-022B-61 | Sequence 61, Appl |
| c 54 | 35.2 | 4.7 | 5008 | 4 | US-07-757-028B-51 | Sequence 51, Appl |
| c 55 | 35.2 | 4.7 | 6211 | 4 | US-08-961-527-8 | Sequence 8, Appl |
| c 56 | 34.8 | 4.6 | 234 | 1 | US-08-469-802B-3 | Sequence 3, Appl |
| c 57 | 34.8 | 4.6 | 234 | 2 | US-08-267-803B-3 | Sequence 3, Appl |
| c 58 | 34.6 | 4.6 | 468 | 3 | US-09-134-001C-2729 | Sequence 2729, Ap |
| c 59 | 34.6 | 4.6 | 854 | 3 | US-08-913-842-16 | Sequence 16, Appl |
| c 60 | 34.4 | 4.6 | 2238 | 2 | US-08-919-624-2 | Sequence 2, Appl |
| c 61 | 34.2 | 4.6 | 1785 | 4 | US-09-468-872-1 | Sequence 1, Appl |
| c 62 | 34.2 | 4.6 | 3964 | 4 | US-08-961-527-136 | Sequence 136, App |
| c 63 | 34 | 4.5 | 198 | 5 | US-08-921-177-6 | Sequence 6, Appl |
| c 64 | 34 | 4.5 | 198 | 5 | PCT-US95-10668-3 | Sequence 3, Appl |
| c 65 | 33.8 | 4.5 | 203 | 5 | US-09-043-303-7 | Sequence 7, Appl |
| c 66 | 33.8 | 4.5 | 397 | 1 | US-08-936-155A-165 | Sequence 165, App |
| c 67 | 33.8 | 4.5 | 8654 | 1 | US-08-920-812-6 | Sequence 6, Appl |
| c 68 | 33.8 | 4.5 | 8654 | 1 | US-08-920-827-6 | Sequence 6, Appl |
| c 69 | 33.8 | 4.5 | 8654 | 1 | US-08-920-828-6 | Sequence 6, Appl |
| c 70 | 33.8 | 4.5 | 8654 | 1 | US-08-921-177-6 | Sequence 6, Appl |
| c 71 | 33.8 | 4.4 | 642 | 1 | US-08-362-577C-6 | Sequence 6, Appl |
| c 72 | 33.6 | 4.5 | 19124 | 1 | US-08-858-207A-48 | Sequence 48, Appl |
| c 73 | 33.6 | 4.5 | 543 | 1 | US-08-487-826B-13 | Patent No. 527390, Ap |
| c 74 | 33.4 | 4.5 | 3280 | 2 | US-08-259-000-4 | Patent No. 5273901-6 |
| c 75 | 33.4 | 4.5 | 3280 | 1 | US-08-729-767-6 | Sequence 6, Appl |
| c 76 | 33.4 | 4.4 | 13440 | 2 | US-09-453-702B-57 | Sequence 57, Appl |
| c 77 | 33.2 | 4.4 | 533 | 6 | US-09-046-578-1 | Sequence 1, Appl |
| c 78 | 33.2 | 4.4 | 1212 | 3 | US-09-285-601-3 | Sequence 3, Appl |
| c 79 | 33.2 | 4.4 | 6709 | 1 | US-09-268-992-7 | Sequence 7, Appl |
| c 80 | 33.2 | 4.4 | 72604 | 4 | US-09-134-001C-462 | Sequence 462, App |
| c 81 | 33.2 | 4.4 | 72604 | 4 | US-09-134-001C-462 | Sequence 462, App |
| c 82 | 33 | 4.4 | 72604 | 4 | US-09-134-001C-462 | Sequence 462, App |
| c 83 | 33 | 4.4 | 1236 | 1 | US-08-257-073-10 | Sequence 10, Appl |
| c 84 | 32.8 | 4.4 | 5181 | 4 | US-09-308-003-6 | Sequence 6, Appl |
| c 85 | 32.8 | 4.4 | 1167 | 2 | US-08-909-965C-11 | Sequence 11, Appl |
| c 86 | 32.8 | 4.4 | 1858 | 1 | US-08-961-527-128 | Sequence 128, App |
| c 87 | 32.8 | 4.4 | 13440 | 2 | US-09-175-658B-22 | Sequence 22, Appl |
| c 88 | 32.6 | 4.4 | 87563 | 1 | US-09-036-574-7 | Sequence 7, Appl |
| c 89 | 32.6 | 4.3 | 198 | 5 | US-09-036-574-7 | Sequence 7, Appl |
| c 90 | 32.6 | 4.3 | 198 | 5 | US-09-036-574-5 | Sequence 5, Appl |
| c 91 | 32.6 | 4.3 | 511 | 4 | US-09-036-574-5 | Sequence 5, Appl |
| c 92 | 32.6 | 4.3 | 652 | 4 | US-08-405-496A-27 | Sequence 27, Appl |
| c 93 | 32.6 | 4.3 | 838 | 2 | US-08-480-604A-27 | Sequence 27, Appl |
| c 94 | 32.6 | 4.3 | 1621 | 1 | US-08-915-136-27 | Sequence 27, Appl |
| c 95 | 32.6 | 4.3 | 3891 | 2 | US-08-961-527-188 | Sequence 188, App |
| c 96 | 32.6 | 4.3 | 3891 | 1 | US-09-734-675-3 | Sequence 3, Appl |
| c 97 | 32.6 | 4.3 | 6492 | 2 | US-08-682-517-27 | Sequence 27, Appl |
| c 98 | 32.6 | 4.3 | 38844 | 4 | US-09-961-527-188 | Sequence 188, App |
| c 99 | 32.4 | 4.3 | 6492 | 2 | US-08-682-517-13 | Sequence 13, Appl |
| c 100 | 32.4 | 4.3 | 3666 | 2 | US-08-682-517-14 | Sequence 14, Appl |

OM nucleic - nucleic search, using sw model

Run on: May 27, 2003, 06:12:17 ; Search time 229.763 Seconds
(without alignments)
4310.289 Million cell updates/sec

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Title: US-09-847-513A-4
Perfect score: 750
Sequence: 1 atgggtaaattattactgat.........ttaaagaatcttctaaatgct 750

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 66023138 residues

Total number of hits satisfying chosen parameters: 1657494

Post-processing: Minimum Match 0%
Maximum Match 100%

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :
```
 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
     Published Applications NA:*
     Listing first 100 summaries
 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 4: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US05_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| c 1 | 44.4 | 5.9 | 6030 | 9 | US-10-239-676-164 | Sequence 164, App |
| c 2 | 44.2 | 5.9 | 14708 | 9 | US-10-239-676-221 | Sequence 221, App |
| c 3 | 43.8 | 5.8 | 1236 | 9 | US-10-077-584-3 | Sequence 3, Appl |
| c 4 | 42.2 | 5.6 | 2614 | 9 | US-09-822-846-491 | Sequence 491, App |
| c 5 | 42.2 | 5.6 | 6604 | 10 | US-09-880-107-1748 | Sequence 1748, Ap |
| c 6 | 42 | 5.6 | 9793 | 9 | US-09-886-156-56 | Sequence 56, Appl |
| c 7 | 42 | 5.6 | 9793 | 9 | US-09-886-149-56 | Sequence 56, Appl |
| c 8 | 42 | 5.6 | 9793 | 9 | US-09-886-150-56 | Sequence 56, Appl |
| c 9 | 42 | 5.6 | 9793 | 9 | US-09-886-159-56 | Sequence 56, Appl |
| c 10 | 39.6 | 5.3 | 5908 | 9 | US-10-239-676-93 | Sequence 93, Appl |
| c 11 | 39.2 | 5.2 | 6730 | 7 | US-08-781-986A-51 | Sequence 51, Appl |
| c 12 | 38.8 | 5.2 | 7906 | 10 | US-10-239-676-97 | Sequence 97, Appl |
| c 13 | 38.4 | 5.1 | 426 | 10 | US-09-960-352-8406 | Sequence 8406, Ap |
| c 14 | 38 | 5.1 | 5986 | 9 | US-10-239-676-145 | Sequence 145, Appl |
| c 15 | 37.8 | 5.0 | 17419 | 9 | US-09-925-302-113 | Sequence 113, Appl |
| c 16 | 37.6 | 5.0 | 2303 | 10 | US-10-239-676-99 | Sequence 99, Appl |
| c 17 | 37.4 | 5.0 | 4239 | 9 | US-09-070-927A-450 | Sequence 450, App |
| c 18 | 37 | 4.9 | 818 | 9 | US-09-764-891-2546 | Sequence 2546, Ap |
| c 19 | 37 | 4.9 | 6053 | 9 | US-09-864-761-2182 | Sequence 2182, Ap |
| c 20 | 36.8 | 4.9 | 2000 | 10 | US-09-887-576-58 | Sequence 58, Appl |
| c 21 | 36.8 | 4.9 | 143306 | 10 | US-09-729-920-3 | Sequence 3, Appl |
| 22 | 36.8 | 4.9 | 1447 | 10 | US-09-983-802-17 | Sequence 17, Appl |
| 23 | 36.6 | 4.9 | 17421 | 9 | US-10-239-676-53 | Sequence 53, Appl |
| 24 | 36.4 | 4.8 | 5979 | 9 | US-10-239-676-17 | Sequence 17, Appl |
| 25 | 36.2 | 4.8 | 12465 | 9 | US-10-239-676-31 | Sequence 31, Appl |
| 26 | 36.2 | 4.8 | 15649 | 9 | US-10-239-676-104 | Sequence 104, App |
| 27 | 36 | 4.8 | 567 | 10 | US-09-070-927A-664 | Sequence 664, App |
| 28 | 36 | 4.8 | 1194 | 10 | US-09-815-242-8601 | Sequence 8601, App |
| 29 | 35.8 | 4.8 | 830 | 9 | US-09-864-761-19531 | Sequence 19531, A |
| 30 | 35.8 | 4.8 | 5898 | 9 | US-10-239-676-179 | Sequence 179, App |
| 31 | 35.6 | 4.8 | 1310 | 9 | US-09-849-243-13 | Sequence 13, Appl |
| 32 | 35.6 | 4.7 | 3263 | 9 | US-09-849-243-15 | Sequence 15, Appl |
| 33 | 35.4 | 4.7 | 4286 | 9 | US-09-849-243-14 | Sequence 14, Appl |
| 34 | 35.4 | 4.7 | 8996 | 9 | US-10-239-676-211 | Sequence 211, App |
| 35 | 35.4 | 4.7 | 425 | 9 | US-09-918-995-8119 | Sequence 8119, App |
| 36 | 35.2 | 4.7 | 11036 | 9 | US-10-239-676-117 | Sequence 117, App |
| 37 | 35.2 | 4.7 | 17142 | 9 | US-09-815-242-9139 | Sequence 9139, App |
| 38 | 35.4 | 4.7 | 17846 | 7 | US-08-781-986A-109 | Sequence 109, App |
| 39 | 35.4 | 4.7 | 488 | 7 | US-09-918-995-19918 | Sequence 19918, A |
| c 40 | 35.2 | 4.7 | 553 | 10 | US-09-920-300A-1461 | Sequence 1461, Ap |
| c 41 | 35.2 | 4.7 | 553 | 12 | US-10-033-528-1461 | Sequence 1461, Ap |
| 42 | 35.2 | 4.7 | 1329 | 11 | US-09-815-242-9139 | Sequence 9139, Ap |
| 43 | 35.2 | 4.7 | 1329 | 10 | US-09-815-242-9604 | Sequence 9604, Ap |
| 44 | 35.2 | 4.7 | 2824 | 12 | US-10-124-557-13 | Sequence 13, Appl |
| 45 | 35.2 | 4.7 | 3066 | 12 | US-10-124-557-83 | Sequence 83, Appl |
| 46 | 35.2 | 4.7 | 3117 | 12 | US-10-124-557-73 | Sequence 73, Appl |
| 47 | 35.2 | 4.7 | 3148 | 12 | US-10-124-557-57 | Sequence 57, Appl |
| 48 | 35.2 | 4.7 | 3420 | 12 | US-10-124-557-103 | Sequence 103, App |
| 49 | 35.2 | 4.7 | 3813 | 12 | US-10-124-557-43 | Sequence 43, Appl |
| 50 | 35.2 | 4.7 | 3936 | 12 | US-10-124-557-141 | Sequence 141, App |
| 51 | 35.2 | 4.7 | 3942 | 12 | US-10-124-557-49 | Sequence 49, Appl |
| 52 | 35.2 | 4.7 | 3945 | 12 | US-10-124-557-45 | Sequence 45, Appl |
| 53 | 35.2 | 4.7 | 3963 | 12 | US-10-124-557-59 | Sequence 59, Appl |
| 54 | 35.2 | 4.7 | 3963 | 12 | US-10-124-557-47 | Sequence 47, Appl |
| 55 | 35.2 | 4.7 | 4086 | 12 | US-10-124-557-39 | Sequence 39, Appl |
| 56 | 35.2 | 4.7 | 4092 | 12 | US-10-124-557-51 | Sequence 51, Appl |
| 57 | 35.2 | 4.7 | 4215 | 12 | US-10-124-557-51 | Sequence 51, Appl |
| 58 | 35.2 | 4.7 | 4575 | 12 | US-10-124-557-57 | Sequence 57, Appl |
| 59 | 35.2 | 4.7 | 5008 | 12 | US-10-044-090-303 | Sequence 303, App |
| 60 | 35.2 | 4.7 | 6665 | 9 | US-10-124-557-1 | Sequence 1, Appl |
| c 61 | 35.2 | 4.7 | 10286 | 9 | US-10-239-676-162 | Sequence 162, App |
| 62 | 35 | 4.7 | 15306 | 9 | US-10-239-676-13 | Sequence 13, Appl |
| 63 | 35 | 4.7 | 2341 | 10 | US-09-881-752A-147 | Sequence 147, App |
| 64 | 35 | 4.7 | 4605 | 10 | US-09-070-927A-315 | Sequence 315, App |
| c 65 | 35 | 4.7 | 7104 | 10 | US-09-815-242-4580 | Sequence 4580, Ap |
| c 66 | 35 | 4.7 | 7107 | 10 | US-09-815-242-8291 | Sequence 8291, Ap |
| c 67 | 35 | 4.7 | 7237 | 10 | US-09-070-927A-239 | Sequence 239, App |
| 68 | 35 | 4.7 | 7434 | 10 | US-09-815-242-4761 | Sequence 4761, Ap |
| 69 | 35 | 4.7 | 7437 | 10 | US-09-815-242-8869 | Sequence 8869, Ap |
| 70 | 35 | 4.7 | 8155 | 9 | US-08-781-986A-63 | Sequence 63, Appl |
| 71 | 35 | 4.7 | 9539 | 9 | US-10-239-676-52 | Sequence 52, Appl |
| c 72 | 35 | 4.7 | 325 | 10 | US-09-878-574-451 | Sequence 451, App |
| c 73 | 34.8 | 4.6 | 449 | 10 | US-09-918-995-26938 | Sequence 26938, A |
| c 74 | 34.8 | 4.6 | 512 | 10 | US-09-864-761-25347 | Sequence 25347, A |
| 75 | 34.8 | 4.6 | 575 | 10 | US-09-864-761-8628 | Sequence 8628, Ap |
| 76 | 34.8 | 4.6 | 6078 | 9 | US-10-239-676-173 | Sequence 173, App |
| 77 | 34.8 | 4.6 | 6203 | 9 | US-10-239-676-183 | Sequence 183, App |
| 78 | 34.8 | 4.6 | 6298 | 9 | US-10-239-676-63 | Sequence 63, Appl |
| 79 | 34.8 | 4.6 | 15732 | 9 | US-10-239-676-96 | Sequence 96, Appl |
| 80 | 34.8 | 4.6 | 319 | 9 | US-09-803-719-2197 | Sequence 2197, Ap |
| 81 | 34.6 | 4.6 | 6620 | 9 | US-10-239-676-195 | Sequence 195, App |
| 82 | 34.6 | 4.6 | 7195 | 9 | US-10-239-676-30 | Sequence 30, Appl |
| c 83 | 34.6 | 4.6 | 465237 | 9 | US-09-933-267A-1 | Sequence 1, Appl |
| 84 | 34.6 | 4.6 | 337 | 10 | US-09-960-352-5976 | Sequence 5976, Ap |
| 85 | 34.4 | 4.6 | 2000 | 9 | US-09-938-842A-2742 | Sequence 2742, Ap |
| 86 | 34.4 | 4.6 | 2000 | 9 | US-09-938-842A-3356 | Sequence 3356, Ap |
| 87 | 34.4 | 4.6 | 3414 | 9 | US-09-074-475-36 | Sequence 36, Appl |
| 88 | 34.4 | 4.6 | 6397 | 9 | US-10-239-676-107 | Sequence 107, App |
| 89 | 34.4 | 4.6 | 20633 | 9 | US-09-815-242-276 | Sequence 276, App |
| c 90 | 34.4 | 4.6 | 293 | 10 | US-09-864-761-18923 | Sequence 18923, A |
| c 91 | 34.2 | 4.6 | 293 | 10 | US-09-864-761-2182 | Sequence 2182, Ap |
| c 92 | 34.2 | 4.6 | 459 | 10 | | Sequence 2182, Ap |